

Grounds for hope on AIDS in Africa

Despite the pervasive tension between AIDS researchers and the South African government, last week's international AIDS conference in Durban was successful on several fronts.

Doubts have been expressed in the past, in *Nature* and elsewhere, about the value of huge, international jamborees of several thousand AIDS researchers and activists, government officials, drug company executives and news reporters. But the 13th International AIDS Conference, which ended last week in Durban, can already claim a number of significant achievements.

One of these has been the way it has focused global attention on the urgent need to make treatment available to people with HIV/AIDS where the epidemic is most serious — in sub-Saharan Africa. Even members of the US Congress normally contemptuous of the concept of foreign aid appear to have discovered the moral imperative of addressing this problem aggressively. Last week the House of Representatives unexpectedly voted for modest additional resources for dealing with the problem.

A second achievement was the meeting's showcasing of several areas of progress in AIDS research that provide hope in the face of the horrifying spread of the epidemic. Researchers reported, for example, that antiretroviral treatments can help to prevent the transmission of HIV from mother to child, even if breastfeeding continues. And preliminary work indicates that the combination therapies now routinely used in the developed world can treat HIV effectively even in poor countries with little health-care infrastructure.

The meeting also brought extra pressure to bear on South African president Thabo Mbeki and his government to address the epidemic with appropriate means and urgency.

The effectiveness of antiretroviral drugs in preventing mother-to-child transmission was a focal point of the conference. Independent studies have shown that short courses of these drugs are both relatively safe and effective. The possibility of transmission through breastfeeding after treatment remains a focus of research, as does drug resistance. But early indications are that, even where breastfeeding

cannot be avoided, transmission of HIV from mother to child is significantly reduced, particularly with early weaning. The meeting also heard evidence of the efficacy of the antiretroviral drug nevirapine, which is cheaper and simpler to administer than AZT.

None of this answers the critical question of how developing countries can afford treatment. But some answers may be forthcoming — for example, last month's offer by five drug companies to provide AIDS drugs to developing countries at reduced prices (see *Nature* 405, 263; 2000), and the offer at the meeting to supply nevirapine free for a limited period (see page 223).

There is no doubt that the urgent need for access to antiretrovirals in developing countries is at last receiving the full attention of the pharmaceutical industry. The South African government has shown little enthusiasm for the offers, but the Durban meeting will help to create international pressure on all parties to reach a compromise that will release the necessary drugs to Africans at affordable cost.

In his closing address to the conference, Nelson Mandela, the former South African president, took an unequivocal stand on the need for his country to implement a programme to reduce mother-to-child-transmission of the virus. His speech was described by conference chairman Jerry Coovadia as a "watershed" in the fractious debate over AIDS that has raged in South Africa since last October.

It is to be hoped that Mandela's words will prompt his successor to reconsider his stand on the aetiology of HIV/AIDS and on antiretroviral drugs. His speech should make it easier for government officials to press for the introduction of treatment programmes for HIV-positive mothers. It will also lift some of the pressure off South Africa's beleaguered physicians and AIDS scientists, who have faced derision and pressure from political leaders over the past nine months. Most importantly, it provides some hope for the tens of millions of Africans who are living with HIV. ■

SOS (again) at the Department of Energy

Congress should provide adequate support for facilities operated on behalf of all US scientists.

It is ironic that, only weeks after politicians so eagerly celebrated the imminent completion of a draft of the human genome sequence, science programmes at the Department of Energy (DoE) should find themselves under pressure again (see page 221).

Ironic, because it was the DoE, with its traditional strength in 'big science', that got the Human Genome Project under way in the first place, in the face of tenacious resistance from the National Institutes of Health (NIH), which later acquired leadership of the project.

The NIH is now enjoying a major build-up of funds that is likely to take its budget above \$20 billion next year. Growing numbers of its investigators rely on large scientific facilities, such as supercomputers or synchrotron light sources, to conduct their

research at the frontiers of molecular biology.

But the NIH doesn't operate any large facilities; they are all maintained by the DoE. It has recently agreed to help equip some of the DoE's synchrotrons, in response to the fact that biologists are among their largest users. But the synchrotrons themselves will continue to rely on the DoE for their operation, maintenance and replacement.

The DoE has many enemies in Washington, but even the most bitter of them expresses no desire to damage the US scientific infrastructure. Sound science policy demands a balance between the support of individual investigators — by the NIH and the National Science Foundation — and the maintenance of major facilities, and their associated research programmes, at DoE laboratories. ■

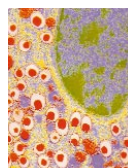




Up, up and away
The first satellites in Cluster II are sent safely on their way
p222



Beating HIV
Mandela calls for urgent action at AIDS meeting
p223



Stem-cell surge
Diabetes therapy could boost demand for stem cells
p224



Up in smoke
Tobacco lobby fails in bid to silence scientific critic
p226

Large research facilities to lose out in science spending spree?

Colin Macilwain, Washington

It is not looking a good year to be a physicist — or even a biologist — who depends on facilities supported by the US Department of Energy. Once again, the department is set to miss out on the large increases in science spending expected to emerge from this year's budget round in Washington.

Scientific societies fear that the many physics subdisciplines sponsored by the department — as well as other scientists who use such facilities — could face a funding crunch this autumn, when the budget process for 2001 is completed.

In February, President Bill Clinton's administration proposed a large budget increase for the Office of Science at the Department of Energy (DoE). He also suggested increases for the other main federal agencies that support basic research, the National Science Foundation (NSF) and the National Institutes of Health (NIH).

Had Congress approved it, the proposal would have raised the Office of Science's funding by 12%, to almost \$3 billion. About half of the increase would have paid for a new Spallation Neutron Source (SNS) at Oak Ridge National Laboratory in Tennessee.

But, as the appropriations process winds its way through Congress, it has become clear that the NIH will get far more than the 5% increase requested by the administration. Lobbyists are confident that the NSF might also get close to the 17% budget boost that Clinton requested.

In contrast, the DoE's science programmes have drawn the short straw. "DoE is the one that is really hurting right now," says David Schutt, head of government affairs at the American Chemical Society.

Last month, the House of Representatives passed a bill that would have reduced construction funding for the SNS from \$280 million to \$100 million, and also cut support for the Basic Energy Sciences and Biological and Environmental Research divisions at the Office of Science.

DoE officials say that the reductions would damage the department's ability to serve the growing demand from NIH-funded biologists for its large facilities.



In demand: biologists are the biggest users of the Brookhaven's synchrotron light source.

At the National Synchrotron Light Source at the Brookhaven National Laboratory in New York, for example, biologists studying protein structure are now the largest user group. "I don't think it is appreciated what an important role DoE now plays in the biological sciences," says Brookhaven's director John Marburger.

Last week, the Senate appropriations subcommittee for Energy and Water, chaired by

Senator Pete Domenici (Republican, New Mexico), passed a plan that would reprove each of these programmes but instead cut the core physics programmes at the agency.

Under the Senate plan, high-energy physics — already struggling to live with Clinton's request (see *Nature* 404, 909; 2000) — would lose \$40 million; nuclear physics and fusion would each lose \$20 million.

The stand-off between these two versions

Jordan chosen to open SESAME

Heather McCabe

The Middle East's international synchrotron research facility, which will bring together scientists from at least 11 countries, has finally been given a home. The new centre will be in Jordan, at a site halfway between the country's capital, Amman, and the West Bank.

The synchrotron has been donated by the German government (see *Nature* 399, 507–508; 1999). Given sufficient funding, it will be rebuilt on property owned by the Al-Balqa' Applied University in Al-Salt.

The site was formally approved last month by the interim council of the project, known as SESAME (Synchrotron-light for Experimental Science and Applications in the Middle East). Armenia, Cyprus, Egypt, Greece, Iran, Israel, Jordan, Morocco, Oman,

the Palestinian Authority and Turkey are already backing the project, and Bahrain, Yemen and Tunisia have expressed their intention to join.

"The project represents a general enthusiasm for collaboration in the Middle East," says Siegbert Raither, director of mathematics, physical and chemical sciences at the United Nations Educational, Scientific and Cultural Organization, which is overseeing the project.

Jordan has agreed to pay US\$1 million per year in operating costs, and member countries are expected to give US\$50,000 per year for the three years of construction.

But more funds are still needed for the project. Organizers are hoping for assistance from the US and European programmes for peace in the region.

of the Energy and Water appropriations bill worries science lobbyists for at least three reasons: the bill is clearly short of money, there is pressure to pass it quickly, and, most importantly, the administration has decided not to veto it.

Domenici would like to sit on the bill until after the August recess, in the hope that it could share in the extra money that Congress and the administration are expected to find in September. But the Congressional leadership wants the Energy and Water bill passed quickly.



Domenici: trying to hold out for extra money.

Mike Lubell, head of government affairs at the American Physical Society, predicts that the House and the Senate will split their differences without providing any extra money — leaving most programmes in the Office of Science facing modest but painful cuts.

Scientific societies are trying to prevent this. But the DoE's science programmes are vulnerable following its security problems with nuclear weapons and fierce clashes between Congressional leaders and its secretary, Bill Richardson.

The Senate has proposed a \$2.7 billion increase for the NIH, and it is thought that, after late budget negotiations, this is what it will get. ■

UK space strategy slammed as lacking funds and vision

Natasha Loder, London

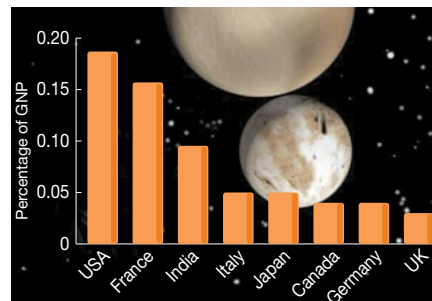
Britain's traditional reluctance to make the same political commitment to space-based activity as its main economic competitors, which has been long criticized by the space community, has now come under strong attack from a cross-party group of Members of Parliament (MPs).

The House of Commons trade and industry select committee has been assessing space policy since last November. In its report, published last week, it says Britain's space strategy is limited in ambition, lacks leadership and is in urgent need of more money — and a radical rethink about how it is spent.

Much of the report focuses on the activities of the British National Space Centre (BNSC), the interdepartmental body set up in 1985 to coordinate the activities of the different parts of government interested in space.

The BNSC is an improvement over previous "ramshackle arrangements" and is successful at a European level, says the MPs' report. But the committee of MPs had received persistent complaints that the agency was failing to coordinate space activities because individual government departments wanted to keep control of their contributions.

The MPs now say the agency should be more proactive, that its role and organiza-



Lagging behind: latest figures put Britain's space spending at the bottom of the world league.

tion should be radically reviewed, and that it should produce annual reports. Also, thought must be given to whether it — or its successor — should have its own budget. Without this, say the MPs, the BNSC's "hands have been tied".

Ian Halliday, chief executive of the Particle Physics and Astronomy Research Council (PPARC), one of the agency's partners, says the BNSC's performance is "effective but not inspirational". This view is shared by Halliday's predecessor Ken Pounds, head of astronomy at Leicester University. Pounds says that for the past 15 years, Britain's attitude to space has been "far too cautious and conservative".

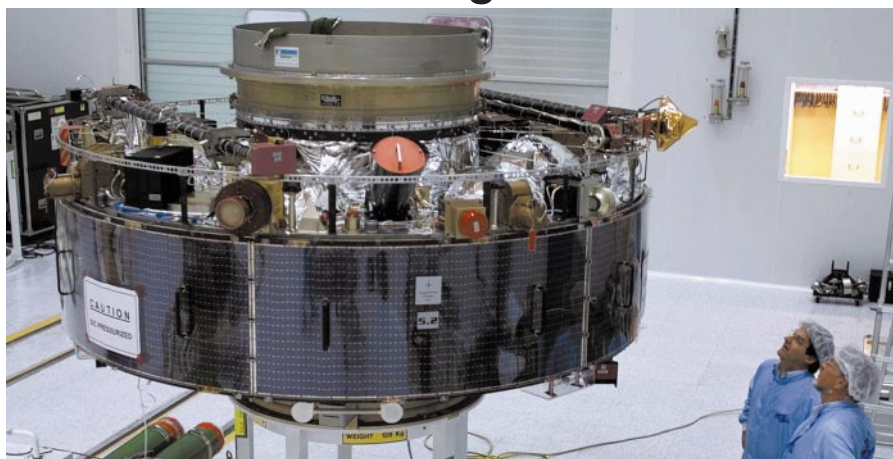
Halliday is also critical of how little money is spent on space science: £30 million of PPARC's contribution to the BNSC goes on membership of the European Space Agency (ESA), and the rest is spent on instruments. Halliday describes ESA membership as comparable to belonging to a golf club while being able to afford only half a set of clubs.

Evidence presented to the select committee shows that Britain's public investment in space R&D is low in absolute and relative terms (see figure). The United Kingdom Industrial Space Committee — the British space industry's trade association — told MPs that "all the main space nations have increased their investment in space in recent years, whereas UK investment has at best remained static".

Although the MPs questioned the proportion of national spending on space compared with contributions to ESA, they also warned strongly against cutting ESA funding. They said this would seriously damage the UK space industry — about 80% of PPARC's contribution to ESA returns to British industry in the form of contracts.

But the MPs do not see greater collaboration with ESA as the best way forward. If the UK space budget were increased, they say, close consideration should be given to whether Britain would do better through bilateral arrangements with NASA, Japan, France, Germany or Sweden. ■

Cluster leaves the ground — at last



In waiting: French engineers inspect one of the Cluster satellites prior to its launch last Sunday.

After a nerve-wracking 24-hour delay, scientists are celebrating the successful launch of two of the four Cluster II satellites that will fly in formation around the Earth, sending back data on the weather in space. They were launched on Sunday at Baikonur Cosmodrome in Kazakhstan, after a last-

minute hitch forced take-off to be postponed. The second pair of satellites will be launched on 9 August. Later this year scientists should start gathering data from the satellites on the structure of the Earth's magnetic field and the flow of charged particles trapped within it. ■

N. L.

Diplomatic Mandela calls for action on HIV...

Michael Cherry, Cape Town

It was not a formal rebuke. But former South African president Nelson Mandela last week came as close as politically possible to chastising his successor, Thabo Mbeki, for re-opening the debate on whether HIV causes AIDS.

Delivering the closing address at the 13th International AIDS Conference in Durban on Sunday, Mandela urged that this debate be "put on the back-burner, so that we can address the needs of those who are suffering and dying".

He added: "History will judge us harshly unless we do so right now." A week earlier, Mbeki's opening address to the conference — in which he refused to respond to demands that he accept the current scientific wisdom on the HIV/AIDS link — had received a hostile reception from many of the 12,500 attending (see *Nature* 406, 113; 2000).

"I do not doubt," said Mandela, "that President Mbeki will proceed with the resolve for which he is known. The challenge is to move from rhetoric to action on an unprecedented scale."

Mandela went on to describe measures to prevent mother-to-child transmission of HIV as "essential", following the examples of Uganda, Senegal and Thailand.

This statement — again an implicit criticism of the South African government's procrastination — was met by tumultuous applause. It also brought a broad smile from conference chairman Jerry Coovadia, professor of paediatrics at the University of Natal, sitting behind the former president on the platform.

"Mandela has raised the debate to a level considered impossible before his arrival at the conference," said Coovadia. So far, the South African government has refused to



A stronger hand: Mandela arrives at the Durban meeting, where he appealed for an end to rhetoric.

introduce an antiretroviral programme for HIV-positive mothers, voicing concern about the toxicity and price of such drugs (see below).

Mandela appears to be the only person in the African National Congress, the main party in the government, with the authority to criticize Mbeki. As such, his comments may prompt a change of heart among cabinet ministers, most of whom have kept silent on the issue.

His remarks will also increase pressure on the government from its own health officials. Many of these are frustrated with Mbeki's stance, arguing that it has held back a commitment to preventing the spread of HIV.

Ayanda Ntsaluba, director-general in the

Department of Health, stated at a press briefing earlier last week that his department would consider implementing an antiretroviral programme for pregnant women as soon as it had received two relevant reports. One of these covers recent trials conducted in the country on the antiretroviral drug nevirapine, the results of which were presented at the conference. The second looks at the obstacles to implementation, based on a study being conducted in the Cape Town township of Khayelitsha.

He stressed that implementation of the programme was not linked to the findings of Mbeki's advisory panel on AIDS, which the president indicated in his address would report back only at the end of the year. ■

... as South Africa considers its options after free drugs offer

Nevirapine's low price and simplicity of administration — it requires only two doses to the pregnant mother, and one to the infant after delivery — has made it a popular antiretroviral drug for preventing mother-to-child transmission of HIV in developing countries.

But the role of pharmaceutical companies remains controversial. On the eve of last week's AIDS conference in Durban (see above) the drug's manufacturer, Boehringer Ingelheim, offered to provide it free to 100 developing countries for five years, as

long as they could demonstrate both that they needed the drug and that they had the infrastructure to administer its provision.

But the offer met with a frosty reception from South African health minister Manto Tshabalala-Msimang, who criticized the company for not telling the government about its offer before the announcement.

Boehringer Ingelheim's representative in South Africa, Lynette Boshoff, told *Nature* that no governments were approached before the offer was made, and that

negotiations would start immediately after the conference with those that were interested in taking it up.

Representatives from Boehringer Ingelheim and four other drug companies and agencies in the Access to Treatment initiative (see *Nature* 405, 263; 2000) met with government officials from across Africa — including Tshabalala-Msimang — in Geneva on 30 June to discuss providing developing countries with cut-price drugs. Boehringer Ingelheim's offer appears to supersede this in the case of nevirapine.

The meeting followed an official approach to the targeted countries, including South Africa, by United Nations agencies to determine whether they wished to investigate the offer. A second meeting is expected to be held in September.

Meanwhile, Tshabalala-Msimang continues to emphasize that South Africa is considering parallel importation (procurement from channels not authorized by the patent holder), and compulsory licensing, although neither option now appears necessary in the case of nevirapine.

M. C.

Diabetes therapy boosts stem-cell campaign

Paul Smaglik, Washington

Recent successes in an experimental treatment for insulin-dependent diabetes are being used by some researchers to support their demands for federal funding of US research using human stem-cells.

The technique, known as the Edmonton Protocol, involves transplanting pancreatic tissue into the livers of diabetics. Devised by a team at the University of Alberta in Canada, the treatment gives the patients healthy islets — groups of cells in the pancreas that regulate blood sugar levels.

According to the researchers, eight patients who have received transplanted islets have not needed insulin therapy for up to 15 months. The results of the experiments will appear in the *New England Journal of Medicine* later this month, but have been released prior to publication because of their potential therapeutic importance.

Widespread use of the technique, however, is likely to require more tissue than can be derived from donors. To achieve therapeutic benefit, the University of Alberta scientists had to obtain islets from an average of two cadavers for every one patient treated.

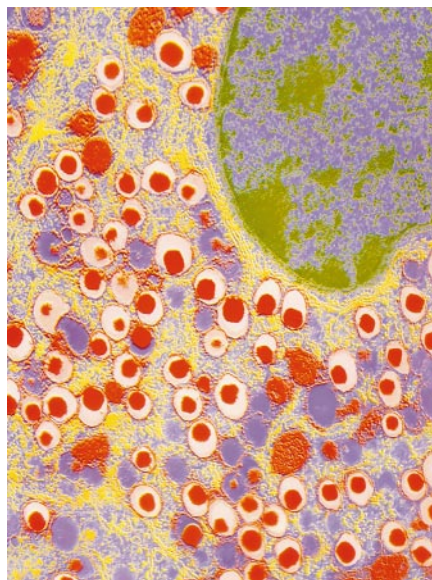
Researchers, including the American Society for Cell Biology, argue that the potential demand for islets bolsters the case for federal funding of stem-cell research. Embryonic stem (ES) cells — if their growth and differentiation can be controlled — might provide an effective substitute to the harvested islets.

Guidelines that would allow the US government to pay for research on ES cells are expected to be agreed by the end of summer. Legislation to expand that funding to the derivation of ES cells has also been proposed (see *Nature* 402, 566; 1999 and 405, 6; 2000).

The main supporters of such legislation, Senators Tom Harkin (Democrat, Iowa) and Arlen Specter (Republican, Pennsylvania), have previously used preliminary reports of the Edmonton Protocol's success to support their proposals for allowing the federal funding of stem-cell research.

Ray Rajotte, a University of Alberta endocrinologist who led the study, says that the effort needed to get islets from conventional sources underlines the need for alternatives. One technique could be to use growth factors or transgenic technology to steer stem cells in the pancreas, rather than ES cells, into becoming the insulin-producing beta cells that form part of the islets.

These 'adult' stem cells are likely to produce only a limited number of beta cells, says Allen Spiegel, director of the US National Institute of Diabetes and Digestive and Kidney Diseases. ES cells, which can divide for a longer period of time, could offer a better alternative.



Growing up? Stem cells may offer an alternative to donor islet tissue (above) for treating diabetes.

But Rajotte warns against overoptimism for both the potential cellular engineering solutions, cautioning that ES and adult stem-cells might not make up for an islet shortage. Although both can be grown into insulin-producing beta cells, he points out that islets also contain alpha cells, which secrete

glucagon, as well as other cell types that, in concert, regulate blood sugar levels.

"We're not sure if the beta cell by itself will work," he says. But Spiegel points to new animal experiments that show beta cells by themselves could correct diabetes.

Previous islet transplantation attempts, which have been largely unsuccessful, used steroids to prevent patients from rejecting organs transplanted along with the islets. But although steroids help to protect the organs from immune responses, they damage the islets.

The Canadian team took several steps to lower the body's immune response. They treated the patients with antibodies to help reduce the production of immune cells, and did not culture the islets before transplantation, to avoid introducing foreign proteins.

An expanded trial of the Edmonton Protocol, funded by the US National Institutes of Health and the Juvenile Diabetes Foundation, will be carried out at ten research institutions in Canada, Switzerland, the United States and Germany over the next 18 months. About 40 patients will receive the treatment, which consists of a 20-minute infusion of the islets through a vein leading directly into the liver. The patients then take daily doses of oral immunosuppressive drugs.

♦ <http://www.med.ualberta.ca/research/groups/islet>

Industry considers carbon catching

Steve Nadis, Boston

While politicians look for economic and social solutions to global warming, industry is intrigued by the possibility of carbon sequestration — preventing carbon dioxide produced by burning fossil fuels from entering the atmosphere.

Earlier this month, the Massachusetts Institute of Technology's (MIT's) Energy Laboratory joined forces with six industrial partners to assess the options for sequestration. These include pumping carbon dioxide from power plants and factories into depleted oil and gas wells and the deep ocean, and planting trees to remove the gas from the atmosphere.

Through the newly formed Carbon Sequestration Initiative (CSI), MIT scientists and engineers will analyse various approaches, while also developing new ideas.

Current sponsors include Norsk Hydro from Norway, Total Fina Elf from France, and the US companies American Electric Power, the Ford Motor Company, General Motors and Texaco. More sponsors are expected. © 2000 Nature Publishing Ltd About \$250,000 anticipated for the first year. The

initial round of research projects will be decided at a sponsors meeting to be held at MIT in late October.

"This is not an effort to implement sequestration ideas," says CSI director Howard Herzog, an engineering researcher at MIT. "We're just trying to determine its potential and identify the key challenges, so that we can make informed decisions when the time comes."

Environmentalists are concerned that sequestration may distract from the goal of reducing dependence on fossil fuels. "Techno-optimism about sequestration could dampen motivation to utilize energy efficiency and renewable energy on the scale we need to," says Peter Frumhoff, a specialist in global change at the Union of Concerned Scientists in Cambridge, Massachusetts.

But Herzog argues that sequestration could be a bridge "that lets us continue using fossil fuels while we develop acceptable alternatives". At the same time, he stresses that any strategies adopted must be environmentally sound. "We want to make sure we're not solving one problem by creating another," he says.

Enhanced advisory role urged for former EU nuclear labs

Quirin Schiermeier

Identifying an appropriate role for the European Commission's Joint Research Centre (JRC) has long been one of the thorniest science-related issues facing the European Union (EU).

Created in 1957 to support the development of nuclear power, the JRC now consists of eight institutes in five countries. But with nuclear power falling from political favour, the centre has increasingly lacked a clear sense of purpose. Today, only 30% of the JRC's Euro1.02 billion (\$950 million) budget is used for nuclear research.

The latest proposal, from an independent panel set up by EU research commissioner Philippe Busquin, is to turn the JRC into a single scientific advice and service institution directly serving the EU's three separate pillars: the European Commission in Brussels, the European Parliament and the Council of Ministers (representing member states).

"The primary function of the JRC should be to facilitate the gathering and fair assessment of information on science and technology matters to inform the EU institutions on a given scientific subject," says the panel, chaired by former industry commissioner Etienne Davignon.

Set up in January, the panel says the JRC should focus on developing monitoring systems or measurement tools in areas relevant to the security of European citizens, including health, food, environment and privacy issues.

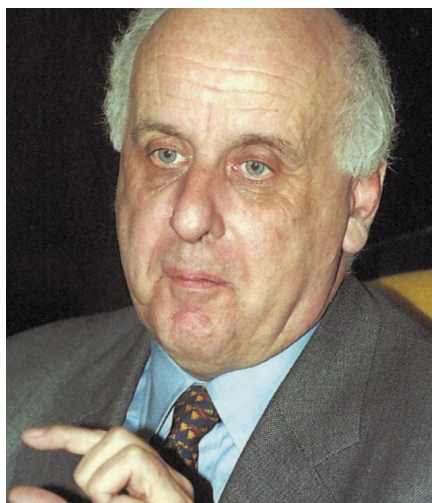
Knowledge gathered at the JRC should be the main source of scientific information for all EU institutions, says the report. But activities should be demand-led: "The JRC should be given a remit that is related to

delivering what the [EU] institutions say they need, not what the JRC thinks they need, as has been the case in the past."

It proposes that the JRC's non-nuclear activities should no longer be funded within the EU's Framework programmes. After 2002, it adds, the JRC's budget should reflect the needs of the three pillars of the EU.

The panel suggests that the Institute for Prospective Technological Studies in Seville, the smallest and youngest JRC institute, be responsible for providing "support and guidance" to the EU's Framework programmes for research, including the development of the sixth Framework programme for 2002–2006.

The report is being seen as a bid to start reform before the arrival of the successor to the JRC's current director Herbert Allgeier, who retires in October. ■



Davignon: wants unified advice for Europe.

Brown boosts British science base

Natasha Loder, London

British science figures high among the beneficiaries of a spending spree announced this week by Gordon Brown, the Chancellor of the Exchequer. Over the next three years, spending on science will average 7% more than this year in real terms, say Whitehall sources.

The figures include a £1 billion (\$1.5 billion) investment in buildings, laboratories and equipment announced last week (*Nature* 406, 113; 2000). But even without this money, £230 million of which comes from the Wellcome Trust, overall spending on research councils and the rest of Britain's 'science base' will be over 4% more than the previous year in each of the next three years.

Although the precise allocation of the

money will not be known until October, £250 million of the increase has been earmarked for research into the advanced computing applications, known as 'the grid', and medicines based on data from the human genome.

The extra money "is clearly very encouraging", says Richard Joyner, dean of research at Nottingham Trent University and chairman of the pressure group Save British Science. But he warns that if university funding continues to decline, there is a danger of an imbalance between the money for new infrastructure and that available to run it.

In his speech, Brown said that in addition to a real increase of 5.4% in the science budget next year, extra resources will be allocated to university-based innovation schemes. ■

Recriminations and confusion over 'fake' coelacanth photo

Heather McCabe, Paris

The hunt is on for the explanation of an apparently faked photograph of a coelacanth. The picture is alleged to be of the fish that French researchers say was discovered in Southwest Java in 1995 — three years before the first official recording of such a 'living fossil' in Indonesia by US researchers (see *Nature* 406, 114; 2000).

Two authors of a recent submission to *Nature* describing the French discovery are claiming innocence and pointing the finger at the third, while admitting that even he may have been duped. But he is also denying any wrong-doing. And the French Institut de Recherche pour le Développement (IRD), for whom all three have worked, has launched an investigation.

The submitted photograph appears to show a coelacanth lying next to three other fish typical of the region. But the image of the coelacanth is identical to a photograph taken by Mark Erdmann of the University of California at Berkeley (see *Nature* 395, 335; 1998).

One of the three authors of the French paper, Georges Serre, a former consultant for IRD's predecessor, ORSTROM, claims to have found the coelacanth in 1995. He says that the specimen he caught was lost on its way to the Indonesian fishery service, and that the photos he took of the fish were stolen.

The two other authors, Bernard Séret, an ichthyologist affiliated with the IRD, and Laurent Pouyaud, an IRD geneticist who works in Jakarta, both claim that Serre either doctored the photo or knew that it was a fake, although Pouyaud also accepts that Serre may have been "manipulated".

Séret sent the photo to an independent expert affiliated with the Tribunal de Paris who asserted that it was a forgery. But Serre — who has told Séret that the photo submitted to *Nature* was taken by a friend who has since died — said last week that he is not yet convinced that the photo is a fake and wants a separate investigation.

Pouyaud also says that he has come across a preserved coelacanth in a private collection near Jakarta, whose owner claims that the fish originated in Java. As it was roughly the same size as the coelacanth that Serre reportedly found, he concluded that it was the same fish. ■

Anti-smoking critic wins campaigning right

Rex Dalton, San Diego

Supporters of the US tobacco industry have failed to silence one of its most prominent academic critics.

An appeals court in California has ruled that Stanton Glantz, a professor of medicine at the University of California at San Francisco (UCSF), can use his position to back the state government's anti-smoking efforts.

In 1997, Californians for Scientific Integrity — a group set up at a Sacramento attorney's office by the industry-sponsored National Smokers Alliance — sued Glantz in an attempt to stop him using his academic position to help the state of California fight smoking.

The group alleged that Glantz's 15 years of campaigning was a misuse of public funds. It also charged that the university's approval of his activities was a violation of the rights of those Californians who did not agree with the anti-smoking cause.

But two weeks ago, a three-judge panel of the Third District Court of Appeals in Sacramento rejected the allegations, reaffirming the right of any faculty member at the University of California to campaign against smoking.

"It makes me very glad to work at a uni-



Last gasp? A Florida jury recently awarded \$145 billion in damages against tobacco companies.

versity that takes academic freedom and public interest very seriously and has the courage to stand up to ... the tobacco industry," said Glantz, who is widely known for challenging the industry's claims that it was unaware of the dangers of smoking.

Citing statements on a recruitment flyer, he claims that Californians for Scientific Integrity "was set up for the express purpose of suing the University of California and silencing me".

"It was clear intimidation," he adds. "But this shows you can stand up and beat them."

Mark Perry, a Washington DC-based attorney with the firm of Gibson, Dunn and Crutcher that represents the suing group, did not answer a request for comment. Perry has said separately that no decision has been reached on whether to appeal to the California Supreme Court.

Nor was there a response from the Claremont Institute, a conservative think-tank in California that filed a 'friend of the court' brief in support of the group.

Although the decision only addresses the rights of university staff to campaign against smoking, it could be far-reaching. Along with other cases, any faculty member of a Californian public university could use it to affirm his or her right to campaign for any cause endorsed by the state legislature.

Last month, Glantz, who conducts research in UCSF's cardiology division, won the annual Public Service Achievement Award of the lobby group Common Cause for his "tireless and tenacious" efforts against smoking. In the past, pro-smoking groups have attacked him in both California and Washington — accusing him of falsifying a result in a study of the effects of 'second-hand smoke', and seeking to cut off his research funds. ■

German librarians cautious of brave new digital world

Quirin Schiermeier, Munich

German librarians are quietly concerned by a recommendation that they increase their efforts to provide electronic services. Although supportive in principle, they are wary of being asked to do too much, too quickly.

The call comes from Germany's science council, the Wissenschaftsrat, which has looked at electronic publishing and archiving in its proposals on the future development of the German science system, published last week (see *Nature* **406**, 116; 2000).

But librarians warn, for example, that there are as yet no satisfactory solutions for reliable long-term archiving of digital media. They also point out that, contrary to many predictions, the emergence of electronic publications has not led to a decrease in print publications.

One of the council's suggestions — an electronic catalogue of the entire stock of German scientific institutes' printed and electronic resources — is already approaching completion.

The Wissenschaftsrat also says that as

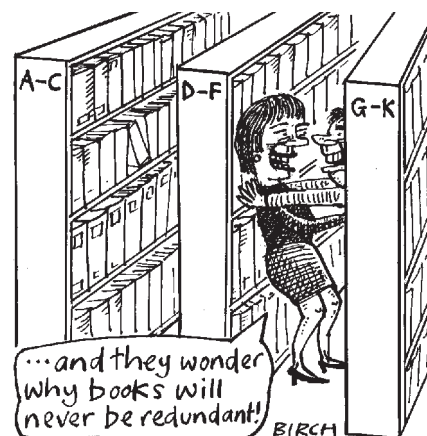
many existing library resources as "economically justifiable" should be electronically converted. Scientists in Germany, it argues, should have remote full-text access to as many publications as possible.

But many librarians believe this is over-optimistic. They argue that the systematic digitizing of library stocks would be of little use for most scientists, and unrealistically expensive. Desktop access, they say, should be limited to key journals and frequently used reference books.

"Digitization projects require a lot of time and money," says Uwe Drewen, head of the working group for scientific librarianship at the Deutsche Forschungsgemeinschaft (DFG), Germany's main research funding agency. "They have to be selected with extreme prudence."

Since 1997, the DFG has approved DM21 million for 63 projects aimed at digitizing selected library stocks. It has also helped to set up digitizing centres in Munich and Göttingen which support libraries in converting print material.

As for direct electronic publishing, the



Wissenschaftsrat accepts that there is no satisfactory solution yet for the long-term archiving of digital publications.

According to Volker Henze, head of the archiving department of the Deutsche Bibliothek (DDB), Germany's national library, no technical solution for guaranteeing the accessibility and integrity of files over the coming generations of computer formats, languages and internet platforms has so far proved feasible.

Meanwhile, the rate at which printed versions of books, journals and university proceedings are deposited in the DDB shows no signs of slowing down. ■

European molecular biology body set to expand its influence

Munich Julio Celis, head of the Institute of Medical Biochemistry at the University of Aarhus in Denmark, has been elected president of the European Molecular Biology Conference (EMBC). Celis takes up his post at a time when the conference, which has previously restricted its activities to funding the European Molecular Biology Organization (EMBO), is starting to flex its muscles.

Earlier this month, the EMBC formally adopted two major projects proposed by EMBO. These are E-Biosci, a proposed electronic repository for primary literature in the life sciences, and the EMBO Young Investigator programme. EMBC now plans to set up a strategic committee to generate more ideas for expanding its portfolio, and hopes to become a strong voice in research commissioner Philippe Busquin's European Research Area, a plan to create a coordinated structure for research in Europe.

Turtles win in latest battle with US fisheries service

Washington Scientists at the US Center for Marine Conservation (CMC) in Washington are celebrating a court victory in their battle to protect the Pacific leatherback turtle. The scientists had sued the National Marine Fisheries Service for failing in its statutory duty to monitor effects of the fast-growing longline fishing industry (see *Nature* 405, 495; 2000).

CMC scientists, with others from Drexel University in Philadelphia, say that the leatherback turtle is on the verge of extinction in the Pacific. Study populations are showing 20–30% annual rates of mortality. A district court in Hawaii has now extended an earlier injunction on longline fishing in the Hawaiian Islands to cover 6.5 million square miles, while the fisheries service has been ordered to produce an environmental impact statement.

Nuclear pioneer who campaigned for peace

Sydney The eminent Australian nuclear physicist Sir Mark Oliphant has died in Canberra at the age of 98. A student at Adelaide University, where the physics department had been established by William Bragg, Oliphant left Australia for England to study under Ernest Rutherford at the University of Cambridge, and worked with him on a number of classic experiments, including the discovery of tritium and helium-3.

During the Second World War, Oliphant left the University of Birmingham to work on the Manhattan Project, building the atomic bomb, but later became a lifelong

campaigner against nuclear weapons. Returning to Australia, he founded the Research School of Physical Sciences at the new Australian National University in Canberra, and became the first president of the Australian Academy of Science.

On retirement, Oliphant became governor of South Australia and a highly respected advocate of the social responsibilities of scientists.

Patent Office seems to have found the secret of success

Munich The number of patent applications filed with the European Patent Office continues to grow. In 1999, the EPO received nearly 122,000 applications, up 7.4% on the previous year. High-tech applications account for most of the rise, with a 15.6% increase in the area of medical technology and a 25.6% increase in electrical communications. In contrast there was only a 3.5% increase in biotechnology.

The EPO is continuing to recruit new staff to cope with the workload. It can afford to do so. Having been embarrassed by its own financial success, the EPO has cut its filing fees by 40% over the past four years — even so, it has still been able to announce a DM250 million (\$120 million) surplus for 1999.

Faulty installation made cable-car crash

Paris The cable-car accident that killed 20 radioastronomy workers in the French Alps last July (see *Nature* 400, 104; 1999) was caused by a mistake made during installation, according to a report by legal investigators. The report, filed with the local magistrate in Gap, adds that the cable-car had frequently been overloaded, contributing to the eventual breakdown in the clamp connecting the cabin and the cable line.

The group was travelling to work at the Institute of Millimetric Radioastronomy's

observatory on the Bure plateau when the cable-car became detached and fell 80 metres. The facility, with five 15-metre telescopes, is operated by France's Centre National de la Recherche Scientifique (CNRS), Germany's Max Planck Society and Spain's National Geographic Institute.

'Safe' Alzheimer's vaccine could reduce damage

Washington A vaccine for Alzheimer's disease that eliminates some of the brain disorder's harmful effects in animals appears to be safe for humans, say researchers. They announced preliminary results of the first human trials at the World Alzheimer's Congress in Washington last week.

Tests in mice showed that the vaccine inhibits the build-up in the brain of amyloid plaques and neurofibrillary tangles, protein deposits characteristic of the disease (see *Nature* 400, 116; 1999).

President Bill Clinton later announced that the National Institutes of Health is to spend \$50 million over the next five years to accelerate research on the disorder.

Harvard subscribes to Celera databases

Washington Harvard University last week subscribed to four of Celera Genomics' databases. The multiyear deal, whose financial terms remain undisclosed, is the second that the biotechnology company, based in Rockville, Maryland, has reached with an academic institution.

Vanderbilt University agreed to subscribe in May. Five large pharmaceutical firms already subscribe to a more detailed version of the company's database and bio-informatics tools. But Celera's president, Craig Venter, has said repeatedly that the company will need academic subscriptions in order to survive.

Kimberly Carr

We regret to announce the death of Kimberly Carr from Hodgkin's lymphoma. Kimberly joined *Nature* in 1992 after obtaining a first degree at the California Institute of Technology and carrying out doctoral research at the French biomedical research agency INSERM.

In her editorial role at *Nature* she was much respected by colleagues and researchers for her skill and judgement in assessing and editing biology manuscripts. She also earned widespread admiration as a writer and broadcaster on scientific topics. Kimberly bore her illness over several years with great fortitude, and will be sorely missed.

Philip Campbell



Bones, molecules...or both?

Evolutionary trees constructed by studying biological molecules often don't resemble those drawn up from morphology. Can the two ever be reconciled, asks Trisha Gura

When biologists talk of the 'evolution wars', they usually mean the ongoing battle for supremacy in American schoolrooms between Darwinists and their creationist opponents. But the phrase could also be applied to a debate that is raging within systematics. On one side stand traditionalists who have built evolutionary trees from decades of work on species' morphological characteristics. On the other lie molecular systematists, who are convinced that comparisons of DNA and other biological molecules are the best way to unravel the secrets of evolutionary history.

Many systematists find themselves somewhere in between these two extremes, accepting that both types of data have their strengths. And lately, some of these moderates have adopted an approach dubbed 'total evidence', in the hope of constructing trees, or phylogenies, to which everyone can subscribe. Rather than considering molecular and morphological evidence separately, total evidence combines both types of data in the same computer analysis. So far, the approach has relatively few converts. Many systematists argue that such disparate types of data cannot readily be combined — and even those who back total evidence disagree over how to analyse the data to avoid one type of evidence dominating the analysis. But if the schism that divides systematics is ever to be bridged, say the adherents of total evidence,

achieving some sort of synthesis of molecules and morphology is essential. "We have to come up with techniques that integrate all sorts of molecular and morphological information," argues Ward Wheeler, curator of invertebrates at the American Museum of Natural History in New York.

The whales' tale

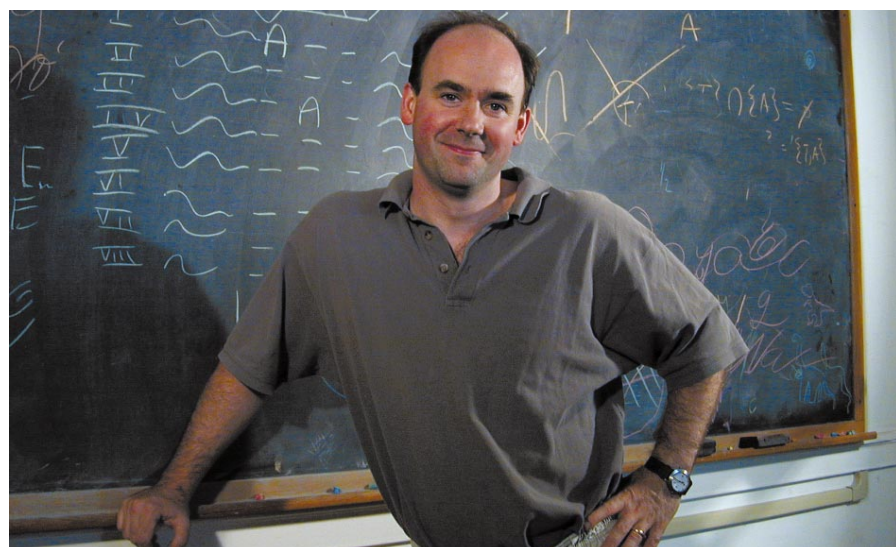
Battles between molecules and morphology are being fought across the entire tree of life. Perhaps the most intense are in vertebrate systematics, where molecular biologists are challenging a tradition that relies on studies of fossil skeletons and the bones and soft tissue of living species. Take the debate over the origin of whales: systematists have long classified even-toed hoofed mammals, called artiodactyls, as a group, or 'clade', that could be further broken into four extant subclades — camels and llamas; cattle and deer; pigs and peccaries; and hippopotamuses. Modern cetaceans — whales, dolphins and porpoises — don't have toes but early fossil whales did have even-numbered appendages. The teeth of ancient cetaceans look almost identical to those from an extinct group called the mesonychids. So morphologists have classified the cetaceans and mesonychids together as a sister group to the Artiodactyla.

But that view came under fire when molecular evidence entered the arena. The first

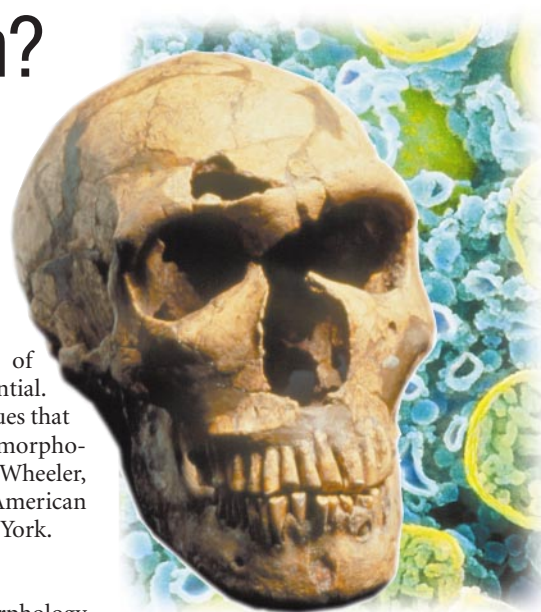
The battles are forcing systematists to adopt more quantitative, reproducible methods.

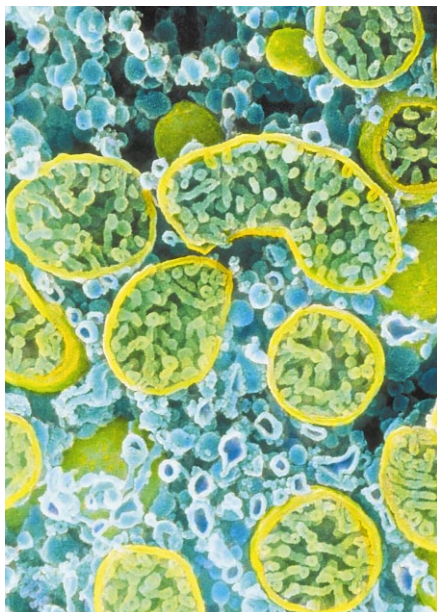
paper to challenge tradition came from Dan Graur and Desmond Higgins at Tel Aviv University in Israel¹. As in many studies of molecular evolution, they focused on the DNA carried in cellular organelles called mitochondria, which placed cetaceans in the middle of the Artiodactyla. Other teams, studying a range of different genes and proteins, singled out hippos as more closely related to whales than to any other artiodactyl, be it a cow, camel or pig²⁻⁵. "Genes doing very different things are all saying the same thing," says John Gatesy, now at the University of Wyoming in Laramie, who was in the vanguard of this molecular assault on taxonomic orthodoxy.

Then came what Graur describes as a "magic bullet": molecular markers that, once present, appear to act as one-way 'signposts' of evolutionary ancestry. In 1997, Norihiro Okada's group at the Tokyo Institute of Technology studied two families of a particular class of retrotransposon — a type of 'jumping gene' — called short interspersed elements (SINEs). These sequences have inserted themselves permanently into random spots in the genome of various mammals and been passed down the generations. After comparing the SINEs found at seven locations in the genome for artiodactyls and whales, the Japanese researchers again concluded that hippos are more like whales than other members of the Artiodactyla⁶. In



Wheeler: if you analyse enough data of many types, convergent evolution should be less of a problem.





Mitochondria: does their DNA link us to Eve?

August 1999, they added further evidence by analysing ten additional genomic locations for the insertion of SINEs and another type of transposon called long interspersed repeats (LINEs)⁷.

Redrawing this branch of the mammalian evolutionary tree according to the molecular data does more than remove hippos from their traditional position, however — it suggests that the Artiodactyla, as currently defined, is not a natural taxonomic group. “The molecular data say that artiodactyls, as traditionally recognized, go into the wastebasket,” says Zhaxi Lou, a palaeontologist at the Carnegie Museum of Natural History in Pittsburgh, who recently reviewed the morphological and molecular evidence in *Nature*⁸. “The whole picture of how we reconstruct the steps of evolution would change,” adds Maureen O’Leary, a palaeontologist at the State University of New York in Stony Brook. “We might be able to say that carnivorous whales were derived from some fully herbivorous animal like ancient hippos.”

Not surprisingly, many researchers —

O’Leary included — are reluctant to simply discard decades of morphological work. “We just don’t know enough yet to determine who is right,” says Hans Thewissen, a palaeontologist and anatomist at Northeastern University College of Medicine in Rootstown, Ohio. He believes the answer may have to await further fossil discoveries. “From the morphological side we need to be smarter about how we pick the animals we analyse and go out and get better fossils.”

If molecular data have provoked strong reactions among researchers interested in the evolution of whales, that is nothing compared to the hornet’s nest stirred up among palaeoanthropologists. The arguments erupted in 1987, when Rebecca Cann, Mark Stoneking and the late Allan Wilson of the University of California at Berkeley introduced the study of mitochondrial (mt) DNA to the field of human origins. Palaeoanthropologists had long asserted that, based on fossil records, modern humans and extinct hominids such as Neanderthals evolved from one common ancestor called *Homo erectus*, which lived in Africa around 2 million years ago. Also, many subscribed to the ‘multiregional’ view of modern human origins, championed by Milford Wolpoff of the University of Michigan in Ann Arbor. This argued that, when the descendants of

H. erectus left Africa, they interbred with hominids in Europe and Asia — including the Neanderthals — giving rise to the ethnic variability seen in modern populations.

Debate rages on

Cann, Stoneking and Wilson’s studies of mtDNA from people of African, European, Asian and Australian descent suggested otherwise. They traced all modern humans to a single female line, dubbed ‘mitochondrial Eve’, which lived in Africa 200,000 years ago, and could find no evidence of interbreeding with other hominids⁹. Although some of the methods and assumptions of Cann and her colleagues have since been questioned, other molecular studies have backed their broad conclusions^{10,11}. Most recently, comparisons of modern human DNA with samples recovered from Neanderthal fossils have failed to find any evidence that modern Europeans carry Neanderthal genes^{12–14}. Still the debate rages on. “We still need better methods that make fewer assumptions,” concludes Cann.

In April this year, Mark Collard at University College London and Bernard Wood at George Washington University in Washington DC published a paper that issued a broader challenge to palaeoanthropology’s traditional morphological approach —



PETER SCOONES/BBC WILD



GEORGE LEPP/CORBIS

Making a splash: molecular analyses shook up evolutionary trees derived from morphology by suggesting that hippos’ closest relatives are whales.

which relies heavily on studies of skulls and teeth. "It struck me that after 15 to 20 years of doing morphological analysis, we hadn't validated the methods we were using," says Wood. So the researchers constructed an evolutionary tree based on 129 skull and tooth measurements for living hominoids, including gorillas, chimpanzees, orang-utans and humans, and did the same with 62 measurements recorded on Old World monkeys, including baboons, mangabeys and macaques. They also drew upon published molecular phylogenies.

At the outset, Wood and Collard assumed the molecular evidence was correct. "There were so many different lines of genetic evidence pointing in one direction," Collard explains. But no matter how the computer analysis was run, the molecular and morphological trees could not be made to match¹⁵ (see figure, below). Collard says this casts grave doubt on the reliability of using morphological evidence to determine the fine details of evolutionary trees for higher primates. "It is saying it is positively misleading," he says. The abstract of the pair's paper stated provocatively that "existing phylogenetic hypotheses about human evolution are unlikely to be reliable".

Many researchers dispute this conclusion, however. "You can't take some 100 arbitrary traits, plug them into a computer, and expect them to agree," says Owen Lovejoy at Kent State University in Ohio.

So can the disparities between molecular and morphological trees ever be resolved? Some proponents of the molecular approach claim there is no need. The solution, they say, is to throw out morphology, and accept their version of the truth. "Our method provides the final conclusion about phylogeny," claims Okada. Shared ancestry means a genetic relationship, the molecular camp argues, so it must be better to analyse DNA and the proteins it encodes, rather than mor-



Family feud: some dispute DNA evidence that Neanderthals (left) did not breed with modern humans.

phological characters that can end up looking similar as a result of convergent evolution in unrelated groups, rather than through common descent. But morphologists respond that convergence can also happen at the molecular level, and note there is a long history of systematists making large claims based on one new form of evidence, only to be proved wrong at a later date.

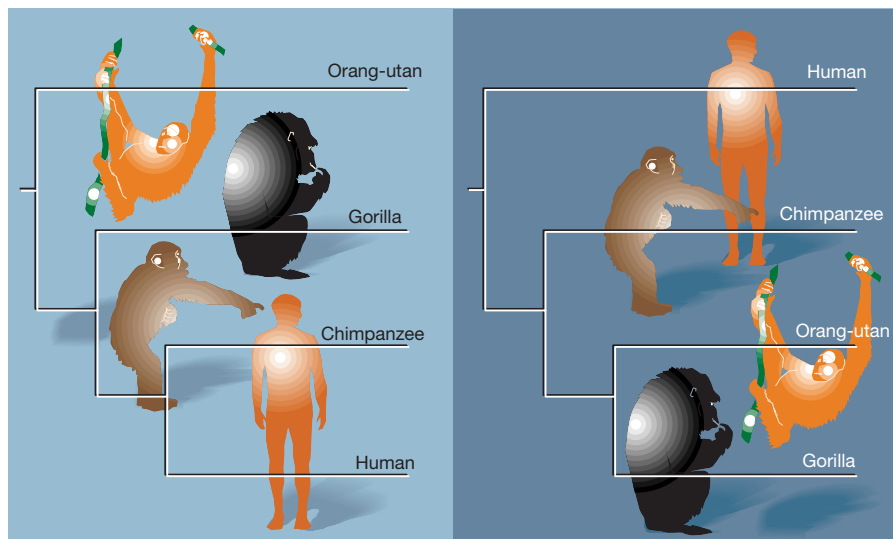
Okada's studies on SINEs and LINEs, held up by the molecular enthusiasts as their

strongest line of evidence, have attracted particular scrutiny. "It is an outdated method in systematics to assert that one aspect of the organism somehow dictates the true phylogeny," says O'Leary. "Okada is approaching this completely backwards by asserting that his retrotransposons are so significant that he cannot imagine a way in which they evolved convergently."

In a commentary published alongside the Okada team's 1999 paper, David Hillis of the University of Texas in Austin pointed out that there is another way in which the retrotransposon evidence could be flawed: regions around the SINE insertion site could mutate so much over time that the DNA elements are no longer recognizable¹⁶. That would obscure links between related groups that diverged in the distant evolutionary past, Hillis argued. "I think we have all become aware that the genetic data have their own limitations and assumptions," says palaeoanthropologist Christopher Stringer at the Natural History Museum in London.

Lovejoy believes progress may come through the study of developmental genetics. For instance, he notes, individual members of the *Hox* gene family can influence the development of several bones. This means that an evolutionary shift towards shorter digits, for instance, could also lead to a shortening of limb length as a by-product, and vice versa. Without knowledge from developmental studies, says Lovejoy, morphologists can mistakenly assume they are dealing with independent characters — and so bias their analysis. "There is nothing wrong with morphology in establishing phylogeny," he concludes, "but it has to be done very carefully with full knowledge of the functional and developmental basis of the traits you are working with." Michael Novacek, a vertebrate palaeontologist at the American Museum of Natural History, agrees that this

Skeletal tissue may be much more plastic in response to subtle changes than was previously thought.



Where do we stand? Molecular (left) and morphological views of relationships among primates.

developmental focus provides “an excellent approach”, but adds that it is limited at present by the availability of data. “The developmental database grows very slowly,” he says.

Unorthodox experiments

For his part, Wood believes that reconciling molecules and morphology may require morphologists to look at different structures. Sally Gibbs, a PhD student in Wood's lab, has combed the literature for non-skeletal anatomical traits that could be used to distinguish five living hominoid genera. In April, at the annual meeting of the American Association of Physical Anthropologists in San Antonio, Texas, Gibbs revealed that a tree based on these characters did resemble the molecular phylogeny. “This is not molecules versus morphology,” concludes Wood. “It seems to be molecules and soft tissue versus hard tissue.”

Wood suggests that skeletal tissue may be much more plastic in response to subtle environmental and genetic changes than was previously thought — a view that finds support from unorthodox experiments by anthropologist Dan Lieberman at George Washington University. He has run pigs on treadmills and analysed their bone content and thickness. Not only did the athletic pigs' leg bones end up stronger and thicker, but so did their skulls¹⁷. “Perhaps growth hormone levels play a role,” Lieberman suggests.

But both soft tissues and molecular data have an obvious limitation: more than 85% of mammalian genera are extinct, and are usually studied from fossils in which bones and teeth are the only structures left. If systematists abandoned skeletal structures they would end up constructing phylogenies around a preponderance of missing creatures.

This is where total evidence, also known as simultaneous analysis, could come to the rescue. First employed in 1989 by Arnold Kluge at the University of Michigan in his phylogenetic studies of *Epicrates*, a genus of snakes in the boa family¹⁸, this approach takes both molecular and morphological data and throws them into the same computer analysis. Although the data may not be consistent, the idea is that misleading evidence will pale in significance as more and more information is added. If you compile enough data, claim the proponents of total evidence, similarities caused by convergent evolution — called homoplasies — will fall out and the evolutionary truth will eventually emerge. “We have an unflappable appetite for data,” says Wheeler, who champions the approach.

In a landmark paper, Kluge and Doug Eernisse of California State University in Fullerton demonstrated the power of the technique by applying it to the phylogeny of the amniotes¹⁹ — a group that embraces mammals and reptiles. Since then, total evidence has been applied to a range of taxa



Wood: realized methods hadn't been validated.

including arthropods^{20–24}, crocodylians²⁵, and primates²⁶. It has also been used to investigate relationships between toothed and baleen whales²⁷.

Suspensions remain

But these attempts at synthesizing molecules and morphology have not yet succeeded in closing the gulf that divides the two camps. At one extreme, molecular biologist Graur rejects any approach that offers an olive branch to the morphologists: “If something is right, why dilute it with something that is wrong?” On the other side, many morphologists remain deeply suspicious of total evidence, seeing it as a means of constructing phylogenies which pay no more than lip service to morphology. Novacek complains that molecular data, which are fast accumulating, will inevitably dominate total evidence analyses. The approach, he argues, may provide a subtle way for molecular biologists to say “Let's get around this traditionalism”.

No one has yet dared apply total evidence to human origins. But the technique has impinged upon the whale–artiodactyl story, albeit so far with inconclusive results. Last year, O'Leary and Gatesy published separate back-to-back total evidence papers examining the problem in the journal *Cladistics*^{28,29}. Gatesy and his colleagues, using datasets on extant species dominated by molecular evidence, backed the whale–hippo link. O'Leary, using a greater amount of morphological data, including fossil evidence, was unable to resolve the data into a single tree. In an attempt to clarify the issue, the pair have now teamed up. Gatesy is sequencing genes from selected whales and artiodactyls, whereas O'Leary is cataloguing morphological features of the same species. They hope to resolve the whale–hippo controversy in one large-scale simultaneous analysis. “If we collect more data from either side, the two signals may come to resemble one another,” says O'Leary.

How the data will be analysed remains to be seen. O'Leary is concerned about the problem of fossils, for which there are morphological data but no molecular evidence.



O'Leary: still has faith in morphology.

She still wonders if it might be best to analyse molecules and morphology separately, to see if the two trees can be made to agree.

Few systematists expect any single study to resolve so vexed a debate, particularly given the strong views on both sides. But most can agree on one thing: the battles between molecules and morphology, and the attempts at synthesizing the two, are forcing systematists to adopt more quantitative, reproducible methods. “There's no more of this ‘one skull here and one skull there,’” says Cann. “You have to have numbers that look like you can do a statistical analysis.” And that, say systematists, is transforming their discipline from a fusty biological sideshow to a vibrant, mainstream field. “There is quite a broad revolution going on as people are being forced to evaluate what they really have,” observes Collard. ■

Trisha Gura is a freelance writer in Cleveland, Ohio.

1. Graur, D. & Higgins, D. G. *Mol. Biol. Evol.* **11**, 357–364 (1994).
2. Irwin, D. M. & Arnason, U. *J. Mamm. Evol.* **2**, 37–55 (1994).
3. Gatesy, J., Hayashi, C., Cronin, M. A. & Arcander, P. *Mol. Biol. Evol.* **13**, 954–963 (1996).
4. Gatesy, J. *Mol. Biol. Evol.* **14**, 537–543 (1997).
5. Ursing, B. & Arnason, U. *Proc. R. Soc. Lond. B* **265**, 2251–2255 (1998).
6. Shimamura, M. *et al. Nature* **388**, 666–670 (1997).
7. Nikaido, M., Rooney, A. & Okada, N. *Proc. Natl Acad. Sci. USA* **96**, 10261–10266 (1999).
8. Luo, Z. *Nature* **404**, 235–237 (2000).
9. Cann, R., Stoneking, M. & Wilson, A. C. *Nature* **325**, 31–36 (1987).
10. Vigilant, L., Stoneking, M., Harpending, H., Hawkes, K. & Wilson, A. C. *Science* **253**, 1503–1507 (1991).
11. Penny, D., Steel, M., Waddell, P. J. & Hendy, M. D. *Mol. Biol. Evol.* **12**, 863–882 (1995).
12. Krings, M. *et al. Cell* **90**, 19–30 (1997).
13. Krings, M., Geisert, H., Schmitz, R. W., Krainitzki, H. & Pääbo, S. *Proc. Natl Acad. Sci. USA* **96**, 5581–5585 (1999).
14. Ovchinnikov, I. V. *et al. Nature* **404**, 490–493 (2000).
15. Collard, M. & Wood, B. *Proc. Natl Acad. Sci. USA* **97**, 5003–5006 (2000).
16. Hillis, D. H. *Proc. Natl Acad. Sci. USA* **96**, 9979–9981 (1999).
17. Lieberman, D. *Am. J. Phys. Anthropol.* **101**, 217–236 (1996).
18. Kluge, A. G. *Syst. Zool.* **38**, 7–25 (1989).
19. Eernisse, D. J. & Kluge, A. G. *Mol. Biol. Evol.* **10**, 1170–1195 (1993).
20. Nixon, K. C. & Carpenter, J. M. *Cladistics* **14**, 221–241 (1996).
21. Whiting, J. M., Carpenter, J. M., Wheeler, Q. D. & Wheeler, W. C. *Syst. Biol.* **46**, 1–68 (1997).
22. Wheeler, W. C. & Hayashi, C. *Cladistics* **14**, 173–192 (1998).
23. Klompen, J. S. H., Black, W. C., Kierans, J. E. & Norris, D. E. *Cladistics* **16**, 70–102 (2000).
24. Baker, R. & De Salle, R. *Syst. Biol.* **46**, 654–673 (1997).
25. Brochu, C. A. *Syst. Biol.* **46**, 479–522 (1997).
26. Horovitz, I. *Am. Mus. Novit.* **3269**, 1–40 (1999).
27. Messenger, S. L. & McGuire, J. A. *Syst. Biol.* **47**, 90–124 (1998).
28. Gatesy, J., O'Grady, P. & Baker, R. H. *Cladistics* **15**, 271–313 (1999).
29. O'Leary, M. A. *Cladistics* **15**, 315–330 (1999).

Caution is sensible, but give biotech a chance to show the good it can do

Sir — On behalf of the Italian Society of Agricultural Genetics (SIGA), I wish to express our astonishment and concern over the televised statement and other reported comments made by Alfonso Pecoraro Scanio, the Italian minister for agricultural and forestry policy, during the International Biotechnology Congress in Genoa on 23–26 May (*Nature* **405**, 388; 2000).

Officials from his ministry were banned from taking part in the congress, despite having previously agreed to present their work on biosafety and on important progress that had been achieved in the ministry's own research institutes. Also, the minister has expressed his intention to ban controlled field experiments of genetically modified plants. In practice, this means a halt to research, whether in applied sectors or in biosafety.

SIGA totally supports the principle that scientific research programmes should undergo rigorous assessment of their objectives, techniques, applications and consequences for society and the environment. Nonetheless, technological innovation is bringing about numerous positive changes, generating increasingly widespread expectations and reactions. There is an uncertain future in store for countries whose state-funded research is under-resourced and inadequate.

SIGA's members (more than half of whom are under 35 years old) are well aware of their responsibilities as academic researchers in the service of the public good — they are not unthinking advocates of the biotechnology market. This is why we feel that the minister's televised invitation to Italian students and researchers to abandon agricultural biotechnology research and devote themselves to the medical sector "which at least saves human lives" is serious cause for concern.

Of course, no human undertaking is risk-free. An innovation is acceptable if its introduction does not prove hazardous by any known criteria. Hence SIGA requests the Italian government — in consultation with the National Commission for Biosafety and Biotechnology, scientific societies and academies, consumer associations, environmental organizations and businesses — to proceed with the application of due precautionary measures to test genetically modified plants in segregated areas, so that their impact on the health and well-being of humankind and the environment can be studied and monitored.

The relationship between science and

society is an inescapable feature of the scientist's profession. Scientists have to question the social effects of their research and appreciate the need for clear communication. The possibility of distorted and hazardous use of new technologies threatens society and causes public mistrust.

On the other hand, scientific research advances the frontiers of knowledge, giving rise to innovations whose production, development and economic use may benefit society. Consequently, biotechnology must be assessed not only in terms of possible risks, but also for the benefits which it may afford in regard to human life and dignity, human rights and values, environmental protection and natural-resource conservation (especially biodiversity).

SIGA is committed to promoting study and research programmes in keeping with these wider goals of science, and to participate in public information programmes.

Enrico Porceddu

*Italian Society of Agricultural Genetics (SIGA),
Department of Agrobiology and Agrochemistry,
University of Tuscia, 01100 Viterbo, Italy*

A fault in the 'weak San Andreas' theory

Sir — Over the past few decades a modern theory of earthquake physics has been developed that is solidly based upon the laws of rock friction¹. Ironically, over that same period there has emerged a view that, alone among crustal faults, California's San Andreas fault grossly violates those laws, while still generating earthquakes indistinguishable from those on other crustal faults. This conclusion is based on the absence of a heat-flow anomaly adjacent to the San Andreas fault that would be expected from a conductive model of frictional heating — a conclusion that requires the friction on the fault to be exceedingly low compared to that of any plausible geological material.

In testing this hypothesis with stress data², I found much of these data ambiguous, but throughout Southern California the data unambiguously indicated that the San Andreas fault is not weak relative to the surrounding crust. This model-independent conclusion is in strong conflict with the heat-flow interpretation. Mark Zoback, in commenting on my results in *News and Views*³, reiterated the heat-flow interpretation as the strongest argument for a low-strength fault, which indeed it is.

As Sherlock Holmes noted about the dog barking in the night-time, the curious thing about the conductive heat-flow anomaly is its absence. There is, however, a broad heat-flow anomaly associated with the San Andreas fault⁴, large enough to account for frictional heating on a strong

fault. Heat flow in this anomaly shows very high spatial scatter, suggesting, together with its breadth, a heat-transport mechanism other than conduction.

Townend and Zoback⁵ show that crustal scale permeability, owing to the presence of active faults, is high enough to ensure that pore pressures are ubiquitously hydrostatic in the crust. Such high fracture permeability would strongly favour advection by fluid transport along fractures, which can transport large heat fluxes without generating massive hot-spring activity⁶. It seems there are cracks in the heat-flow edifice.

C. H. Scholz

Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York 10964, USA

1. Scholz, C. H. *Nature* **391**, 37 (1998).
2. Scholz, C. H. *Geology* **28**, 163–166 (2000).
3. Zoback, M. *Nature* **405**, 31 (2000).
4. Lachenbruch, A. H. & Sass, J. H. *J. Geophys. Res.* **85**, 6185–6222 (1980).
5. Townend, J. & Zoback, M. D. *Geology* **28**, 399–403 (2000).
6. Manga, M. *Geology* **26**, 799–802 (1998).

That famous double helix takes a sinister turn

Sir — The double-helical structure of DNA is one of the most recognizable scientific images, and one of the most frequently abused. In particular, it is often shown as a left-handed instead of a right-handed helix.

The announcement last month that the first draft of the human genome is complete has featured prominently in the national news, frequently accompanied by left-handed helices.

More worrying than this mistake made by the popular press is the frequency with which DNA is wrongly depicted in the scientific press, particularly that of the genetics community. Such errors do little to inspire confidence.

Nature has no excuse for depicting left-handed helices in a figure to an Insight article about functional genomics¹, or for claiming in a Book review that Franklin's X-ray crystallograph "led Watson and Crick to deduce the left-handed double-helical structure of DNA"².

On the contrary, Watson and Crick's original article³ describes (and illustrates!) a right-handed sense.

Colin Porter

*Institute of Cancer Research, Fulham Road,
London SW3 6JB, UK*

1. Vukmirovic, O. G. & Tilghman, S. M. *Nature* **405**, 821 (2000).
2. *Nature* **405**, 737 (2000).
3. Watson, J. D. & Crick, F. H. C. *Nature* **171**, 737 (1953).

Nature replies — This is correct, and we will be more vigilant in future. See <http://www.lmmb.ncifcrf.gov/~toms/LeftHanded.DNA.html> for a comprehensive general discussion of this problem.

The biotechnologist's jungle book

A look at the pharmaceutical treasure trove of the Amazonian rainforest.

Medicine Quest: In Search of Nature's Healing Secrets

by Mark J. Plotkin

Viking: 2000. 304 pp. \$22.95

Walter Gratzer

"If you are fat, fleshy, brightly coloured and slow moving, I want to see you," says Dr David Newman of the National Cancer Institute, and he is not alluding here to the matriarchs of our affluent society. "There's something in you, on you, or travelling with you," he continues, "that stops you from being eaten." Or, as Mark Plotkin puts it in this sprightly, engrossing and often startling account of nature's bounteous therapeutic gifts to man, "the creature with the best chemistry set wins". So, indeed, do ailing humans and, of course, the pharmaceutical companies, whose chemists toil at isolating the channel and receptor blockers, vasodilators, anticoagulants, antibiotics and anaesthetics that these animals and plants so copiously harbour. For, as Plotkin reveals, the US retail take from anticoagulants and thrombolytic agents alone passed \$700 million in 1993 and is rising fast. Morphine and its relatives gross \$600 million per annum, and Plotkin estimates the national cost of pain, measured by drugs, doctors' bills and lost income, at about \$100 billion.

Medicine Quest, then, is no mere collection of arcana of natural history. Some 70% of pharmaceuticals now in use are or derive from natural products. (Most of the rest probably sprang from fortuitous discoveries: physiological effects sometimes came to light when organic chemists cautiously tasted their reaction product. And the anti-cancer drug cisplatin, for instance, resulted from an attempt to assess whether an electric current could affect the growth of bacteria: the misshapen forms that appeared in an *Escherichia coli* suspension resulted, it eventually turned out, from the activity of a reaction product of chloride ions with the platinum electrode.)

Plotkin observes that the majority of the species to which our planet is host live in the seas, and their biology remains largely unexplored. Carnivorous marine snails alone have produced a myriad of physiologically active substances, and it was a sponge that yielded the adenine nucleoside of which AZT is a synthetic derivative. Small wonder, then, even if only one in 10,000 therapeutically promising substances makes it to the market-place, that a shoal of small companies with names such as Cal Bio Marine, Pharma Mar and Oceanix Biosciences have taken to the water.

But Plotkin's real passion is for the Amazonian rainforest and its denizens. Here is the arrow-poison frog; this secretes vasodilators and analgesics, as well as a cardiac-channel blocker that acts on a different calcium pump from that inhibited by digitalis. Then there is the leech, which injects an anticoagulant, a vasodilator and an anaesthetic to render its victim unaware of a six-inch proboscis lustily probing its flesh. The giant, 18-inch leech of nightmares is armed not only with anticoagulants but also with a clot-dissolving enzyme. A vast range of anticoagulants, acting at different levels in the clotting cascade, is to be found in ticks and vampire bats (draculin), and a component of viper venom, which reduces platelet adhesion, is being tested as a treatment for angina. There is the monster toad, which exudes hallucinogens (and was introduced into Australia, where it has established itself as a pest and apparently serves as a lollipop for the local junkies in search of a cheap fix).

Plotkin conjectures that autochthonous tribes learned about the medicinal properties of the indigenous flora by observing the behaviour of animals. A monkey with diarrhoea will peel and chew the bitter bark of the *Vernonia* bush, which African tribespeople use to treat the same condition; pigs

use pomegranate roots, which contain a vermifuge, to treat (apparently) parasitic infestations; tapirs eat the nekoe plant, the toxins of which pass through the intestines; the sagacious animals then, Plotkin assures us, hasten to the nearest river where they deposit their droppings to narcotize the fish on which they like to feast.

Plotkin is plainly entranced by the rainforest, and likes to consort with shamans, who feature prominently in his earlier book, *Tales of a Shaman's Apprentice* (Penguin, 1994). He often shares the life of the forest tribesmen and is not above displaying the occasional flash of bravado ("I grabbed my machete and followed him ... together we entered the twilight-green world of the forest"). He appears at times a little too persuaded of the shamans' powers, and some of his anecdotes of seemingly miraculous cures will stretch the credulity of sceptical readers.

But these are small cavils about what I believe is an important book. Plotkin seeks not only to entertain and inform (in which he abundantly succeeds), but also to warn, especially against some of the consequences of man's destruction and pollution of the world. Francis Bacon declared that nature, to be commanded, must be obeyed. His injunction, as we all know, has not been

Healing hues

Foxgloves are the source of digitalis, a substance with powerful effects on the heart, used to treat congestive heart failure and cardiac arrhythmias.

Nature's Medicine: Plants That Heal by Joel L. Swardlow with photographs by Lynn Johnson (National Geographic, \$35, £20), from which this picture is taken, traces man's search for healing plants through the ages, from the ancient civilizations of China and India through to the birth of modern pharmaceutical science.



book reviews

observed. There are many other reasons why what little remains of the tropical forests should be cherished, but what Plotkin laments here is that the annihilation of species by the thousand deprives us of untold opportunities to discover new medicines and other valuable classes of compounds — strong materials, process chemicals and many more.

He also denounces the profligacy with which existing therapeutic resources are being degraded: antibiotics are still fed to cattle and federal law is content to allow trace amounts of no fewer than 80 antibiotics to enter the milk that Americans drink. And finally, Plotkin notes, a large proportion of the most important medicinal substances were brought to us by research driven only by curiosity about the natural world and mocked by Senator William Proxmire and his allies as a waste of the taxpayers' money. Let us hope that Plotkin's voice is heard in the quarters that count, those in which economic and not moral interests prevail. Dr Johnson put it pithily: "Sir, I have found you an argument, but I am not obliged to find you an understanding."

Walter Gratzer is at the Randall Centre for Molecular Mechanisms of Cell Function, New Hunt's House, Guy's Campus, St Thomas Street, London SE1 1UL, UK.

A function for dysfunction

Ion Channels and Disease

by Frances M. Ashcroft
Academic: 1999. 352 pp. £46.95, \$75

Thomas J. Jentsch

The importance of many things in our lives becomes apparent only when we lose them, or when they malfunction. In genetic disease, mutations in a gene may lead to interesting and sometimes unexpected insights into the physiological importance of their product. Detailed comparison of the properties of both the normal and mutant gene products can reveal deep insights into the physiological role of the protein encoded by the gene and how it goes awry in disease.

This combination frequently occurs in ion-channel diseases (or 'channelopathies'). Over the past 15 years, genes encoding ion channels have been isolated at ever-increasing rates. Their number exceeds estimates obtained from biophysical studies, and one might wonder why there are so many different channels. In many cases, mutations in ion-channel genes, in both humans and animals, have provided clues or answers. Some ion-channel genes were even cloned on the basis of the effects of their disruption. For example, positional cloning in the fruitfly

Channelling other energies

The Kongo peoples inhabit the River Congo region in west-central Africa, and are divided into many subgroups, all of which have many different cults. Their strong belief in a wide spectrum of spiritual forces catalysed the development of a wealth of sculptural forms dedicated to containing and harnessing these forces. The statue shown here, called *nkisi nkonde*, represents a power figure generally associated with vengeance and aggression, but it could also be involved in the sealing of pacts between individuals — each nail or point driven into the statue reinforces the pact. Failure to keep the pact would provoke retribution from *nkisi nkonde's* particularly ferocious spirit. From *Spirits Embodied: Art of the Congo* by Evan M. Maurer and Niangi Batulukisi (University of Minnesota Press, \$34.95, £24.50).



Drosophila led to the isolation of the *shaker* potassium channel, and the CFTR chloride channel was found because its gene is mutated in cystic fibrosis.

The large number of ion-channel diseases provides an impressive illustration of the many functions of channels. These range from signal transduction to ionic homeostasis, cell-volume regulation, trans-epithelial transport, vesicle acidification and trafficking, and cytotoxic effects, to name just a few. As a result, ion-channel diseases are clinically very diverse, ranging from cystic fibrosis to inherited forms of kidney stones, periodic paralysis, high blood pressure, epilepsy and hyperinsulinaemia.

Ion channels can be studied in great detail using biophysical methods. One such method is the patch-clamp technique, which investigates the properties of single molecules in real time. The *in vitro* analysis of channel mutations found in patients has contributed considerably to our understanding of the relationships between channel structure and function. Compared with most other genetic diseases, this provides an unrivalled level of understanding of how mutations cause disease.

Frances Ashcroft's book is a fascinating and accessible account of ion-channel

diseases. But it does not stop there—it is also an up-to-date description of the different classes of cloned ion channels, their structure and function, and how the disease characteristics resulting from their dysfunction are used to elucidate their physiological roles. For completeness, channels not (yet) known to cause human disease are also covered, and the chapter "Ion channels as lethal agents" even discusses the pore-forming peptides of bacteria and fungi.

Ashcroft begins with a short introduction on the essentials of molecular biology, human genetics, electrophysiology and ion-channel design. Although necessarily superficial, this section will be useful for students from different backgrounds. Beginning with voltage-dependent sodium channels, the author goes on to discuss the different classes of ion channels individually. Only cloned ion channels are included, thereby omitting several classes of ion channels (such as calcium-activated chloride channels) whose molecular correlate has not been firmly established.

These chapters begin with a concise description of what is known about the structure and function of the channels, and a discussion of the key experiments that led to these insights. The physiological role of the

channels is then described. These short sections introduce the reader to the cellular physiology of such diverse systems as the β -cells of the pancreas, photoreceptors, kidney tubules, the neuromuscular junction and even the complement system.

Ashcroft then discusses diseases that are caused by mutations in these channels. The main focus is on human disease, but mouse models of disease, channel mutations in goats and horses, and in *Drosophila* and the roundworm *Caenorhabditis elegans* are also covered. The perspective is broadened by the inclusion of such aspects as the modulation of inositol phosphate metabolism by lithium in the treatment of manic depression, and a discussion of how glutamate-receptor ion channels are affected by the ingestion of excitotoxins (from shellfish, Nutrasweet or glutamate in the Chinese restaurant syndrome).

Not only are these sections interesting, they are also fun to read. We learn what to pay attention to when preparing a meal from the pufferfish *Fugu*, or how to differentiate clinically between myotonia and paramyotonia (give the patient ice-cream!). Each chapter is illustrated with useful figures and diagrams, and remaining problems are highlighted. For instance, the pathophysiological mechanisms underlying some important channelopathies (such as cystic fibrosis) are still not completely understood many years after the gene was identified.

Ion Channels and Disease is a very good introduction to the whole field of ion channels, and will be useful for researchers, clinicians and students. Although Bertil Hille's *Ionic Channels of Excitable Membranes* (Sinauer, 1992) remains the standard textbook on the biophysical aspects of ion channels, and *Ion Channels* (Cambridge University Press, 1996) by David Aidley and Peter Stanfield is a very good, balanced overview of the structure and function of such channels, Ashcroft's book combines a short review of individual ion-channel classes with an excellent discussion of their physiological and pathophysiological roles. It is the favourite of several of my colleagues. Although the book is as up-to-date as it can reasonably be, I hope that frequent new editions will keep pace with this rapidly developing and fascinating area. ■

Thomas J. Jentsch is in the Zentrum für Molekulare Neurobiologie, Universität Hamburg, Falkenried 94, D-20246 Hamburg, Germany.

More molecular biology

Biology of Sensory Systems

by C. U. M. Smith

Wiley, £39.95, \$80 (pbk)

Molecular and Cellular Physiology of Neurons

by Gordon L. Fain

Harvard University Press, \$65, £40.50



Feathered flagships: kagus, endemic to the biodiversity hotspot New Caledonia.

Conservation gets heavy

Hotspots: Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions

by Russell A. Mittermeier, Norman Myers & Cristina Goettsch Mittermeier
University of Chicago Press: 2000. 432 pp.
\$65, £45.50

Rory Howlett

Beautifully illustrated, *Hotspots* represents the second volume resulting from a three-way collaboration between the non-profit organization Conservation International, the Mexican conservation organization Agrupación Sierra Madre and, believe it or not, Cemex, the world's third-largest cement company. Judging from the book's weight, one would be forgiven for musing that Cemex contributed more than just corporate support to its production. That aside, the company's environmental credentials are now no doubt set in concrete.

The first volume in the series, *Megadiversity: Earth's Biologically Wealthiest Nations*, published in 1997, dealt with the so-called B-17, the 17 biologically richest nations. The present volume focuses on 'biodiversity hotspots'. The hotspots concept, conceived of and promulgated by ecologist Norman Myers, is nebulous, but, roughly speaking, highlights the importance of those regions of

the world that have disproportionately high levels of species richness and, in particular, endemism.

Defining the thresholds of biodiversity and endemism above which a region qualifies for 'hotspot' status is arbitrary. Indeed, as the first chapter makes clear, hotspots have been subject to 'reanalysis and reassessment', with areas being added or removed as criteria for what constitutes a hotspot have evolved. The basic message, though, is that an estimated 60% to 70% of all terrestrial plants and non-fish vertebrates occur within just 1.44% of the Earth's land surface, providing a rationale for what Myers has advocated as a 'silver bullet' approach to biological conservation.

The following 25 chapters document separate terrestrial ecoregions, each chapter authored by specialists. The organization is formulaic. Overviews of the regions, their diversity of habitats and global importance are followed by subheaded sections on 'Flagship species', 'Threats' and 'Conservation'. This makes it easy to access information, despite the absence of an index. For example, the chapter on New Caledonia starts by pointing out that, although the region has a land surface area roughly equivalent only to that of Wales or New Jersey, it is remarkably species rich, with high levels of endemism. In this it is like Madagascar, once also part of the ancient continent of Gondwana. Flagship taxa include a rainforest-dwelling bird, the

book reviews

kagu (*Rhynochetos jubatus*), the largest arboreal pigeon (*Ducula goliath*) and a great diversity of endemic geckos and skinks. Environmentally unfriendly open-cast nickel mining is a significant threat, as are forest logging and the introduction of alien species. The section on conservation documents the often inadequate protection status of the unique ecotypes of the region, and lists recommendations for the short term; these include such platitudes as "Development of legislation and policies that promote biodiversity conservation and sustainable resource management".

Peter Seligmann, chairman of the board of Conservation International, contends in his preface that this is an "important book". Whether, in time, it will be seen as such will depend on whether it informs the minds and touches the hearts of the power-mongers. Perhaps more likely, the volume will adorn living-room coffee tables and be the subject of anguished after-dinner conversation about the state of the environment, as well-heeled guests admire the glorious colour prints and sip their Brazilian coffee. ■

Rory Howlett is deputy biological sciences editor at Nature.

Chemistry's Canterbury Tales

A Chemical History Tour: Picturing Chemistry from Alchemy to Modern Molecular Science

by Arthur Greenberg
Wiley-Interscience: 2000. 312 pp.
\$59.95, £38.95

John Emsley

People in Britain agonize over joining the European monetary system that will be based on a new currency, the euro. And yet 2,000 years ago there was only one currency in Europe and the Middle East, that of the Roman Empire. The gold solidus and silver denarius were legal tender from the north of England to the south of Egypt, from the east of Turkey to the west of Africa.

Political expediency meant that the coinage was often debased, but the Roman emperor Diocletian, who ruled from AD 285 to 300, was determined to stabilize the currency, and, as part of his financial reforms, he ordered the destruction of all alchemical texts. He feared that the practitioners of the black art might find a way of transmuting lead into gold and undo his good work.

Consequently, all the alchemical works throughout the empire were consigned to the flames and with them went the chemical knowledge that had been acquired over the centuries in ancient Egypt. Had it not been for Diocletian's misguided directive, Arthur

Greenberg's delightful book might well have contained even more fascinating accounts of the early days of chemistry. Even so, he has written 120 short essays based on books and papers that span the history of chemistry from the Ancient Greeks to Nobel prizewinners.

Greenberg has put his collection in chronological order, grouping them into eight sections: mining of metals; alchemy (including the tricks used to fool the gullible); early medical treatments; the emergence of chemistry as a science; the birth of modern chemistry; farming and industry; teaching and popularizing; and, finally, the theories of bonding. No matter where you dip into this book you will find something of interest, although an item in the alchemy section on the colour of the pottery of the Catawba Indians of North Carolina seems somewhat irrelevant.

Otherwise, the reader can be assured of an informative and entertaining read. The section entitled "Chemistry begins to emerge as a science" is a treasure trove, dealing as it does with the works of Johann Glauber (after whom Glauber's salt, or sodium sulphate, is named), Robert Boyle (regarded as the founder of chemistry), Georg Stahl (he of the phlogiston theory), Stephan Hales (long forgotten, but the inventor of apparatus for dealing with gases), Henry Cavendish (who discovered hydrogen and who unknowingly isolated the noble gases) and Joseph Priestley (of oxygen fame).

My own interest in popularizing chemistry caused me to turn first to the section, "Teaching chemistry to the masses", where I discovered the works of Mrs Jane Marcet, whose 1805 best seller *Conversations on Chemistry* (160,000 copies eventually sold) awoke the interest of the young Michael Faraday. Not only did he turn out to be one of the greatest scientists who ever lived, but he acknowledged his debt to her influence by becoming an equally good popularizer of chemistry himself.

Nor does Greenberg shy away from the more embarrassing attempts by Victorians to give chemistry a popular spin. There is mention of Lucy Rider Meyer's fanciful *The Fairy*

Land of Chemistry, in which each element is portrayed as a fairy, and molecules become charming groups of interacting sprites. Greenberg also quotes from J. Carrington Sellars' poetic *Chemistianity*, written in 1873, as if chemistry really were a kind of religion. In his piece entitled "And now turn to page 3 of our chemical psalm book", Greenberg quotes an example of its achingly bad verse:

OXYGEN, the Queen of Body Affection;
The supporter of man's Earthual life;
The needed Air-puff for all common forms
Of combustion in term'd live Animals.

Not surprisingly, *Chemistianity* appears to have made few converts. Yet it is such occasional items that make this book so attractive. Of course, most of the book deals with the seminal ideas in the emergence of chemistry — the story of the discovery of oxygen and the downfall of the phlogiston theory are particularly well told — but its real value lies in the way it acts as a chemical *Reader's Digest*, capturing the flavour of books that are of historical import, but which the average chemist would rarely see or, indeed, now want to read.

Chemistry has a rich heritage of which it can be proud, and yet along the highway to our present state of knowledge there were many by-ways and cul-de-sacs to be explored. The rich tableaux of characters encountered in *A Chemical History Tour* is worthy of *The Canterbury Tales*, and the reproductions of paintings, illustrations from rare old texts and Greenberg's whimsical asides make it particularly enjoyable. Nor is it just an Old Testament of chemistry. Greenberg takes us right up to the modern day in the final section, where he deals with the origins of stereochemistry, radiochemistry, atomic and molecular structure and DNA, ending with the scanning tunnelling microscope and the realization that we can now study atoms one by one. ■

John Emsley is science writer in residence in the Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK.

New in paperback

Mapping the Deep: The Extraordinary Story of Ocean Science

by Robert Kunzig
Sort of Books, £8.99

"I value this book for its almost exclusive attention to the neglected deep oceans. If it is read by enough lay people, perhaps I shan't have to explain so often that, no, oceanographers don't all go to sea wearing red watch caps, and we don't all have scuba tanks slung over our shoulders." Alan Longhurst, *Nature* 399, 425–426 (1999) — from the review of

The Restless Sea, the title under which the book was originally published in 1999.

Empires of Time: Calendars, Clocks and Cultures

by Anthony Aveni
Tauris Parke, £11.95

Rosalyn Yalow: Her Life and Work in Medicine

by Eugene Straus
Perseus Books, £11, \$16

Jumbo

The revolutionary machine that shrank the world.

Vaclav Smil

The Boeing 747 was the first wide-bodied long-range jet — but this terse description does not do justice to a plane which most flight aficionados see as the best jetliner ever built. As it takes off for an intercontinental flight that will last more than 12 hours, a fully loaded 747-400, the latest in the jumbo series, weighs close to 400 tonnes, the equivalent of more than six of the stubby, short-range 737-500s. About 40% of its mass is jet fuel, and its four engines, mounted on wings spanning more than 64 metres, deliver a peak combined thrust of almost 112,000 kg, equal to about 280 MW of power.

The Boeing 747 was developed at the urging of Juan Trippe, founder and then president of Pan American World Airways (Pan Am). It was neither the first airliner powered by gas turbines nor the first successful intercontinental jet: the ill-fated British 106 Comet 1 presaged the jet age in 1952, and it began in earnest in 1958, with the redesigned 106 Comet 4 and Boeing 707.

William Allen, Boeing's boss, gambled the company's future by investing more than twice its worth to build the world's largest passenger jet. The wide body (to allow for two standard ship containers to be placed side-by-side) and the bubble cockpit (for easy loading through an upturned nose) betray the ultimate design intent: in the late 1960s it was expected that supersonic jets would soon

dominate all long routes, and 747s would become cargo carriers. That never happened — but the 747, rather than being just another compromise, turned out to be, in the engineer Henk Tennekes' memorable description, the only plane that obeys ruthless engineering logic.

To accomplish this, Boeing had to overcome a string of design and financial crises brought about by the ambitious construction schedule, and Pratt & Whitney had to develop the first giant commercial turbofan engines. Incredibly, the prototype took off on 9 February 1969, less than three years after Pan Am ordered its 25 747s. The first scheduled flight on 21 January 1970 came back to the gate after an engine overheated during taxiing; more than five hours later a back-up plane finally left New York for London.

If only the company's troubles had ended there! During the past three decades, Boeing has seen massive cyclical layoffs, strikes, productivity problems, plummeting stock prices and, most recently, a fierce competition with Airbus, whose smaller and slower A340 has a longer range than the 747, and

Boeing's ambitious construction schedule brought about a string of financial and design crises.

whose planned super-jumbo A3XX is to carry 550–650 people. And Pan Am are not even around any more.

But the accomplishment remains. There are successful companies that produce little of lasting value; and there are troubled, or defunct, enterprises whose daring and vision have changed the world. The Boeing 747 is an icon and a favourite of the jet age. In 1988 came the first redesigned version, the 747-400: passengers do not see the less cluttered digital avionics and programmable displays in the two-crew cockpit, but they can feel the power of new high-bypass turbofan engines, they can see the cruising speed on their monitors closing on 0.9 Mach, and as many as 524 of them can fly nonstop at the edge of the stratosphere from Vancouver to Hong Kong, or from Singapore to London.

The 1,000th 747 was delivered on 12 October 1993, and the count passed 1,230 at the beginning of 2000. So far, the Queen of the Skies — or, more affectionately, the Big Top or Fat Albert — has carried more than two billion passengers, equal to nearly 40% of humanity. Her size, range, speed and economy have made intercontinental flight routine, with steadily declining real costs, as tens of millions of people go to family reunions, business meetings or beaches, and billions of dollars of electronic goods, flowers, clothes and toys move to rapidly changing markets. The growing integration of the global economy in general, and Asia's economic rise in particular, have been closely tied to the international fleet of 747s.

The Boeing 747 combines symbolism and function, beauty and economy: a daring, revolutionary design that has become a powerful symbol of global civilization, an improbably graceful behemoth that ushered in the age of mass intercontinental travel, its allure has not faded after three decades of unprecedented service. Next time you are at a large international airport, find a gate with a 747. Then stand, childlike, with your face against the glass right in front of the plane's huge rounded nose, note the four massive four-wheel bogies, look up at the bubble topping the fuselage, check the enormous engines mounted on a gently arched, variably cambered wing, and sweep to its end where an upturned winglet makes it easy to identify the most modern 747-400 series. Like a Gothic cathedral or a high-flying flock of Canada geese, the 747 is an object of undeniable beauty, graceful daring and irresistible awe. How can that thing fly? How that thing can fly! ■

Vaclav Smil is at the University of Manitoba, Winnipeg, Manitoba R3T 2N2, Canada.

MARK WAGNER/AVIATION IMAGES

All is not lost

By resurrecting extinct infodiversity, we may save our own culture.

Scott Westerfeld

The discovery of the Landry-M'batu 'Cretan Spirals' could substantiate theories about a global proto-language from which all modern languages descend. The successful decoding of the Spirals also represents another addition to the recent spate of dramatic reconstructions of palaeo-data. Whether or not the implications for panhuman language hold up, the Spirals certainly provide cheer for the generally pessimistic new science of global infometrics.

The Spirals are found on a class of Cretan pottery from about 3,800 years before the present (around the sixteenth century BCE). These 'running spirals' were produced by turning the pottery on a fast wheel while sliding a metal stylus down the side, inscribing a continuous spiral groove. The pots were then glazed, which helped preserve surface features such as the running spiral. Landry and M'batu hypothesized that any words spoken by the potter during the inscription process would vibrate the stylus, and thus be recorded in microscopic undulations of the groove, and that these sounds could be recovered by molecular mapping. After comparing analyses of several pots, the researchers discovered a consistent pattern of recorded vibrations.

The apparent habit of Cretan potters was to utter a standard prayer during inscription of the spirals. The benchmark of this (yet untranslated) prayer's structure has allowed the researchers to compensate for variations in wheel speed and other factors. Ultimately,

the phonemes of the prayer and other utterances were decoded, providing audible reconstructions of Cretan speech. Comparisons with reverse-engineered Sino-Tibetan and other proto-languages are claimed to have fallen in line with the predictions of panhuman proto-language proponents. However controversial these assertions, more heartening to global infometrists was the ability of Landry and M'batu to exhume shreds of an extinct spoken language, a legacy thought buried irrecoverably by time.

Much has been written about the mass extinctions at the beginning of the third millennium. Indeed, in the two-century period of the Great Die-Off, the world lost approximately 65% of its species, with tropical rainforest biomes particularly hard hit. But global infometrics posits that this biotic die-off was merely one aspect of a generalized large-scale information loss that marked the end of the millennium, the so-called Infometric Bottleneck. In addition to biotic depletion, the world lost 85% of its spoken languages; saw the end of all trade systems outside the Common Currency Unit; and the consolidation of world culture to four great religions, a single legal code and a universal software platform. Despite the common complaint that the contemporary world is too complicated, it is by any infometric measure a place of arctic simplicity.

Of course, the neck of the bottle was economic globalization, which, although it created the prosperity that 95% of humanity now enjoys, also flattened differentiation via the wholesale razing of local biomes, social

practices, and belief systems. Despite the engineering of new species, the expansion of language through emergent slang, and the cultural creolization afforded by new communications technologies, a few hundred years is simply not enough time to recreate the lost diversity that existed pre-Bottleneck.

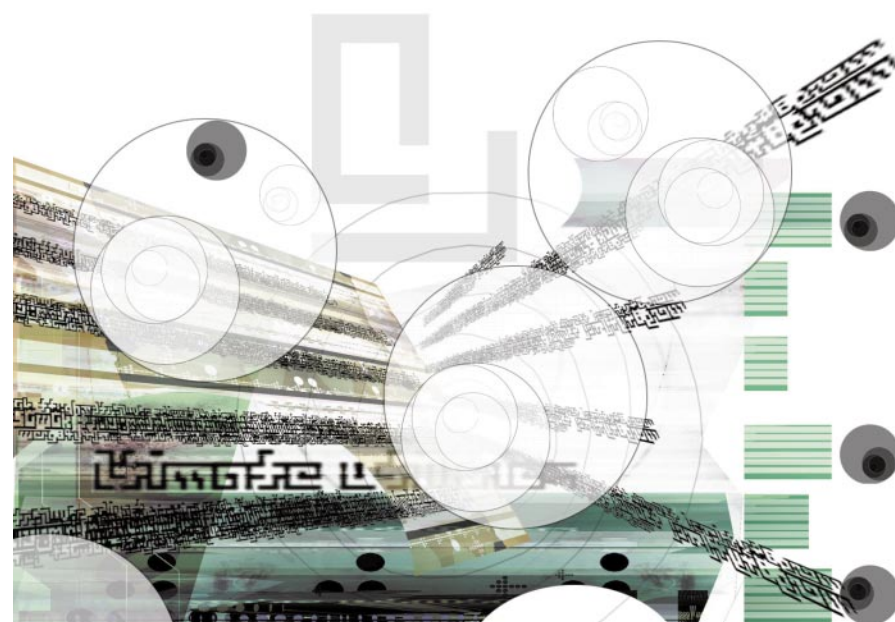
As some recent events have shown (the currency-software crash of 2279, the Australian Biotic Collapse, the Elvis Heresy and so on), this depauperated state of the globe's information reservoirs can have perilous consequences. Simple systems are inherently less stable than diverse ones. The question is, can we reach into the past to retrieve lost information, and replenish our reserves?

New technological innovations seem to suggest that we can. The minute undulations recorded in the Landry-M'batu Spirals could not have been transcribed without molecular mapping, nor interpreted without cohomonym topology, a branch of linguistochastics developed only in the past few decades. The gene canon has allowed recreation of extinct species: the Siberian tiger, the blue whale, a host of red algae. Countless rainforest species are preserved *in vitro* and remain to be reconstructed, if the complex environmental relationships among them can be worked out. The dinosaurs have famously eluded reanimation. The lack of anything like a dinosaur womb in the present has left the codes in recovered DNA undecipherable. Apparently, some information is irrecoverably lost.

Space technologies have also recovered pre-Bottleneck information. Translight probes recently exploring the HD141569 star system recorded faint FM transmissions that left Earth in the 1980s. (Moving at a sluggish lightspeed, these transmissions are only now reaching that star.) Perhaps some decades hence, several concentric spheres of translight craft could be erected at various distances from Earth, and capture the bounty of the relativistic information wake that surrounds the planet. From these transmissions pre-Bottleneck languages, music and customs may one day be reconstructed routinely.

Much like the words recorded on the Landry-M'batu Spirals, these ancient voices would call to us faintly and in unfamiliar tongues, almost swallowed by the noise of their long journey through time. But surely we will listen carefully, their message the more precious for almost having been lost for ever.

Scott Westerfeld's latest novel is *Evolution's Darling*, from *Four Walls Eight Windows* (www.4W8W.com).



Faster than a speeding photon

Jon Marangos

The textbooks say nothing can travel faster than light, not even light itself. New experiments show that this is no longer true, raising questions about the maximum speed at which we can send information.

Can a light pulse travel faster than the speed of light? This question has intrigued physicists for many years because such an event could violate Einstein's theory of special relativity and the principle of causality (that 'cause' always precedes 'effect'). Together these imply that no object or information can travel faster than the speed of light, $c = 3 \times 10^8 \text{ m s}^{-1}$. For nearly two decades, physicists have been sending certain light pulses faster than c over short distances (so-called superluminal propagation), but the light pulses have always been distorted in the process so interpreting these experiments has been difficult¹⁻³.

In May this year, Mugnai *et al.*⁴ reported superluminal behaviour in the propagation of microwaves (centimetre wavelengths) over much longer distances (tens of centimetres) at a speed 7% faster than c . A report by Wang *et al.*⁵ (page 277 of this issue) now demonstrates a very large superluminal effect for pulses of visible light, in which a pulse propagates in a specially prepared medium with a negative velocity of $-c/310$: that is, not only faster than a pulse travelling in a vacuum, but so fast that the peak of the pulse exits the medium before it enters it!

A negative velocity can be understood by comparing the times it would take for identical pulses of light to cover some distance L in a vacuum (travelling at velocity c) and in a superluminal medium (travelling at velocity v). The difference in transit times $\Delta T = L/v - L/c$ is a negative quantity if the velocity is superluminal. If v has a negative value then ΔT can become sufficiently negative that the peak of the pulse emerges from the medium at an instant earlier than when the peak of the pulse enters. This brings to mind Arthur Buller's well-known limerick with relativistic overtones:

*There was a young lady named Bright,
Whose speed was far faster than light;
She set out one day,
In a relative way,
And returned home the previous night.*

But Wang *et al.*⁵ claim that, unlike the heroine of this rhyme, their light pulses do not violate causality. They argue that their superluminal pulses are the result of the wave nature of light itself (fortunately, making it impossible for an object with mass to travel faster than c) and that no actual information, or signal, is transmitted faster

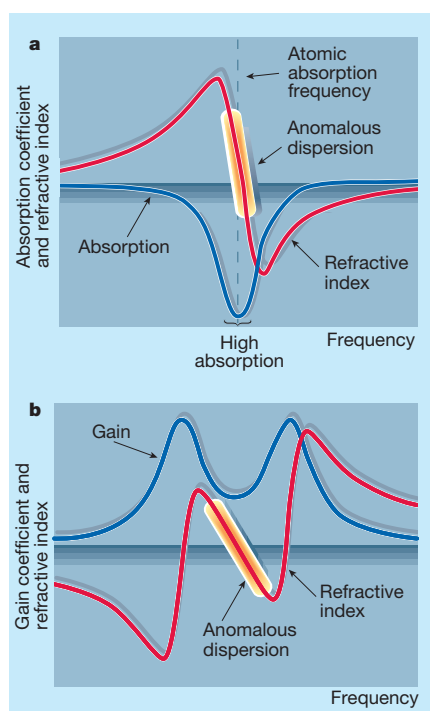


Figure 1 Sending photons faster than light.

a, How light absorption and refractive index of a dispersive material change rapidly with wavelength when the wavelength of the light pulse is near an atomic absorption band. The anomalous-dispersion region (where the group velocity of light can be negative) coincides with a region of strong light absorption. b, How the gain and refractive index of caesium gas changes with wavelength when there is a 'gain doublet' (two closely spaced peaks) in the amplification of light. Wang *et al.*⁵ show that in this case the anomalous-dispersion region can be used to make a pulse of light travel faster than c .

than c . They use smooth, well-defined light pulses, so that the peak of the pulse at the output results from the forward rising edge of the input pulse, which occurs far earlier in time, making it consistent with causality. An abrupt feature in the light pulse would not be able to travel faster than c . This means that even if the 'effect' appears to precede the 'cause', you still can't send useful information — such as news of an impending accident — faster than c .

A light pulse has a finite duration, and it is a well-known theorem in physics (the bandwidth theorem) that, to create a pulse

of finite duration, an infinite number of waves of different frequency must be added together. The shorter the desired pulse, the larger the bandwidth of frequencies that must be used. All light pulses are therefore formed by a packet of waves of different frequency, each of which has a different amplitude and phase. There is a distinction between the speed of individual waves, called the phase velocity, v_p , and the velocity at which the peak of the wavepacket propagates, known as the group velocity, v_g . In a vacuum the phase and group velocities are the same, but in a highly absorbing or dispersive medium they are usually different. A negative group velocity results when the phases of the different frequency components are shifted by the medium through which they travel, so that the wavepacket they form at the exit is brought forward in time compared with the same pulse travelling through a vacuum.

One way to achieve negative velocity is to modify the refractive index of the medium through which the light passes. Last year scientists at Harvard⁶ and elsewhere succeeded in modifying the refractive properties of a cloud of ultracold atoms to generate very slow light pulses with group velocities of a few metres per second. To create the opposite effect — superluminal pulses of light — you need a medium in which the refractive index changes rapidly, for example near an atomic absorption frequency (Fig. 1a). The only problem is that the so-called anomalous dispersion region in Fig. 1a, where v_g can be negative³, is also in a region where there is increased light absorption. In experiments with such highly absorbing materials, the light pulses are either strongly distorted or absorbed, making any faster-than-light claims difficult to interpret.

A more promising approach to making superluminal light pulses is to work with an atomic medium where there is gain (amplification of light waves) at the atomic transition frequency. This is achieved in a laser-type medium by creating a 'population inversion', whereby a higher population of atoms are in the excited than in the lower-energy atomic state⁷. In this case, anomalous dispersion occurs at frequencies lower than the transition frequency. But close to the transition frequency, where the effect is largest, the rapidly changing gradient in the refractive index causes severe pulse distortion. One

way round this problem is to use a gain doublet⁸ — that is, two closely spaced regions of gain where the zone between has steep anomalous dispersion but without strong pulse distortion (Fig. 1b). This is what Wang *et al.*⁵ have now achieved.

The experiment by Wang and co-workers creates this type of gain doublet in a six-centimetre cell containing caesium gas by using two laser fields closely spaced in frequency (see Fig. 1a on page 277). They first measured the refractive index of the caesium using a third 'probe' laser, and produced a dispersion curve similar to Fig. 1b, with a steep gradient in the anomalous dispersion region corresponding to an expected $v_g = -c/330$. When they sent a 3.7-microsecond light pulse through the medium, it appeared at the exit of the cell before it arrived at the entrance. Although the pulse itself is only shifted forward in time by a modest fraction (1.7%) of its width, this corresponds to the wavepacket leaving the cell 62 nanoseconds before it arrives — in other words, travelling nearly 20 metres away from the cell before the incoming pulse enters it. Compared with the time to travel six centimetres in a vacuum (about 0.2 nanoseconds), the 62-nanosecond lead means that the group velocity of the pulse inside the medium is $-c/310$, close to the predicted value.

In this experiment, each of the different frequency components making up the pulse experiences a slightly different dispersion in the medium. The relative phases between them are therefore changed and the pulse shape is shifted to bring the pulse wavepacket (or group velocity) forward in time. So the anomalous dispersion leads to interference between different frequency components of the pulse that produce the superluminal effect. Although amazing, this type of superluminal pulse propagation does not violate the principle of causality.

There remains, however, some debate about what is the true speed at which information is carried by a light pulse. Traditionally the signal velocity of a light pulse is defined as the speed at which the half peak-intensity point on the rising edge of the waveform travels; in this experiment, this is clearly superluminal. In contrast, some researchers argue that the true speed at which information is carried by a light pulse is not the group velocity of a smooth pulse, but rather the speed at which a sudden step-like feature in the waveform travels, which so far has not been shown to exceed c . Superluminal effects are especially interesting in the case of light pulses consisting of only a few photons, in which it could be argued that the group velocity is the same as the velocity of the individual photons. The type of superluminal behaviour discussed here is also predicted to apply to single photons⁸, which might have implications for the transmission of quantum information. ■

Jon Marangos is in the Blackett Laboratory, Imperial College, Prince Consort Road, London SW7 2BZ, UK.

e-mail: j.marangos@ic.ac.uk

1. Steinberg, A. M., Kwiat, P. G. & Chiao, R. Y. *Phys. Rev. Lett.* **71**, 708–711 (1993).
2. Balcou, Ph. & Dutriax, L. *Phys. Rev. Lett.* **78**, 851–854 (1997).
3. Chu, S. & Wong, S. *Phys. Rev. Lett.* **48**, 738–741 (1982).

4. Mugnai, D., Ranfagni, A. & Ruggeri, R. *Phys. Rev. Lett.* **84**, 4830–4833 (2000).
5. Wang, L. J., Kuzmich, A. & Dogariu, A. *Nature* **406**, 277–279 (2000).
6. Hau, L. V., Harris, S. E., Dutton, Z. & Behroozi, C. H. *Nature* **397**, 594–598 (1999).
7. Chiao, R. Y. *Phys. Rev. A* **48**, R34–R38 (1993).
8. Steinberg, A. M. & Chiao, R. Y. *Phys. Rev. A* **49**, 2071–2075 (1994).

Biophysics

Science in motion

Philip Ball

If it moves, it's biology', goes the saying, but the moment one asks how it moves, physics intervenes. Between the release of chemical energy and the buzz of a fly's wing there is a host of minor mechanical miracles. At the other end of the scale, co-operative motions of groups of organisms ranging from bacteria to fish and humans can influence the movements of individuals. Can it be a coincidence that both fish and slime mould cells swarm in the same patterns (Fig. 1)? A workshop* in Budapest last month showed how physical and biological sciences can collaborate to explain these miracles of motion.

Motion in biology is a cascade of mechanical transduction processes, from the molecular scales of time and length upwards. At every level in this hierarchy, biological motion provides perhaps the ideal testing ground for the current migration of physicists towards biological problems. At the molecular level, the two have long been blended in conventional biophysics, which has been revitalized by single-molecule probe techniques. It is hard to imagine how, without these, one could unravel the secrets of muscle proteins such as titin — the spring that gives relaxed muscle its elasticity. For example, the inequivalence of stretching and contracting in individual titin molecules can be attributed to rapid unfolding and slow

refolding of repetitive protein domains (M. Kellermayer, Pécs Univ., Hungary).

But whereas the movements of motor proteins like myosin and kinesin and springs like titin are being decoded residue by residue, it appears that something else may be needed to convert protein movements into the motion of whole cells. Migrating cells are central to embryo development and wound healing.

Myosin typically collects at the trailing edge of the cell to pull it in the direction of motion, whereas a polymerizing network of actin pushes the leading edge forward (G. Borisy, Univ. Wisconsin-Madison). How can actin, a relatively limp filamentary polymer, develop any force to push on the cell membrane? The answer, it seems, is that repeated branching of the polymers, at an angle close to 70°, ensures a constant supply of short filaments at the front edge of the network. These are stiff enough to generate the necessary force. The branching is initiated on the inside of the membrane itself by a membrane-bound protein called Wasp, which activates a second protein, Arp 2/3, to secure itself to actin and provide a junction for a branching filament.

But how do new actin monomers add to the advancing tip if it is pushing against the membrane? Here help is on hand from physics, specifically from George Oster's idea

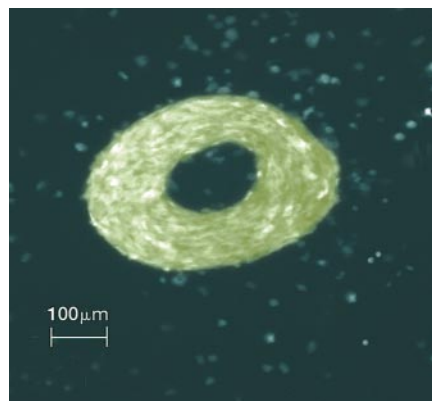
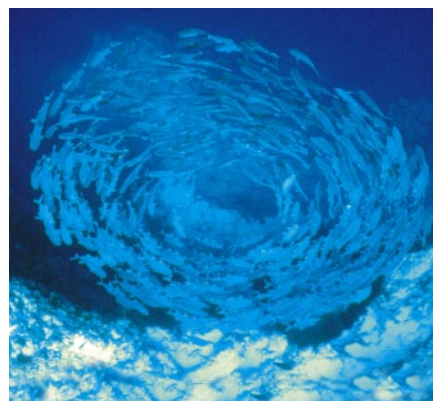


Figure 1 Swirling vortex motion is a mode of collective swarming behaviour exhibited by both fish (left) and slime mould cells (right).

of a brownian ratchet. Thermal fluctuations of the flexible actin filaments expose the tip long enough for a new monomer to attach. Once the tip springs back, it ratchets the membrane forward by the length of one monomer. Brownian ratchets have found a life of their own in the physics literature; it is good to know this amounts to more than idle curiosity.

Granted that cells move, how do they know where to go? Under some circumstances — for example, during chick embryogenesis — tracking specific cells labelled with multicoloured fluorescent proteins seems to indicate a more or less random walk (R. Lansford, California Inst. Technol.). But *in vitro* studies of nerve cells known as astrocytes seem to show directional streaming motions into unpopulated areas (A. Czirók, Eötvös Univ., Budapest). M. Abercrombie suggested 50 years ago that mutual adhesion was enough to guarantee directional motion at the leading edge of a tissue. But there is directionality behind this edge too, which demands a more complex model.

Commonly, cell migration is guided by trails of attractant and repellent chemicals: the process of chemotaxis. Single-celled organisms apply this trick too, finding safety in numbers in times of stress. The best studied of these systems is the slime mould *Dictyostelium discoideum*, which aggregates into a multicellular 'slug' — a kind of 'super-organism' — when food is short. The slug cells differentiate into two types: one forms a long stalk, the other a 'fruiting body' containing spores, which will survive almost indefinitely until revived by favourable conditions. Virtually all stages of this process can now be modelled by making some simple assumptions about cell-to-cell interactions (H. Levine, Univ. California, San Diego).

Following migration, the spontaneous sorting of young cells into regions of different cell type can create coherent structures in tissues. This may be nothing more than a variation of the kind of phase separation observed between immiscible fluids (G. Forgács, Clarkson Univ., Potsdam; J. Glazier, Univ. Notre Dame). Differential adhesion between the various cell types, which is mediated by adhesion molecules at the cell surface, may create something analogous to surface tension for a cell cluster, and energy minimization would then take care of the rest. But, as ever in biology, one cannot take it for granted that variables such as surface tension will not change over time in an active cell.

At the ecosystem level, the 'why' of aggregation behaviour is not so obviously answered as for embryo cells or slime moulds. For animals that form groups, the benefits may depend on a delicate balance between such factors as foraging efficiency, distribution of the spoils and chances of attracting predators (J. Parrish, Univ. Washington). The diversity of spatial patterns arising from the tendency

to aggregate offers rich grounds for physical modelling, and is as yet little understood.

Human behaviour offers another layer of complication: the veneer of social conventions. In social sciences the tradition has been to assume that these dominate — that patterns of motion are dictated primarily by considerations such as courtesy. Hence the challenge posed by simulations in which people are represented by particles moving under little more compunction than that for a particular velocity and for collision avoidance (D. Helbing, Dresden Univ. Technol.). It is unnerving to watch these streams of disk-shaped pedestrians spontaneously arrange themselves into counter-flowing streams in a corridor, or passing in groups through a door before standing back and giving a chance to those coming in the other direction. And most chilling of all is to watch these automata converge on a narrow bottleneck under conditions of mass panic,

and collectively block the exit in the crush to escape.

There are useful lessons for both parties from this kind of meeting between biology and physics. Even the most conscientious of physicists cannot expect to get all the necessary information from a textbook, nor to be confident without input from biologists that the neglected details are really dispensable or the resemblances between model and experiment more than skin deep. The biologists can learn that approximations are permissible even in the most complex of systems, and that complex behaviour can (even if that does not mean it must) arise from simple principles. As one speaker remarked, to make the interaction work, "we all have to be in the same room."

Philip Ball is a Consultant Editor at Nature.

*Models of Biological Motion, Collegium Budapest, Institute for Advanced Study, Hungary, 19–22 June 2000.

Neurobiology

Parallel sensing

Mathew E. Diamond

Having a poorly developed visual system, rats use their whiskers to navigate, to explore, to detect objects, and to determine the size, shape and texture of those objects¹. The 30 or so whiskers on each side of the snout sweep back and forth about four to ten times a second. This 'whisking' motion greatly expands the space that can be sampled, but it also complicates the calculations that the brain has

to perform to identify the location of an object. The information carried by the nerves emanating from the whiskers is complex: what was touched, and where and when it was touched, must all be represented in the brain. This could easily overload the capacity of a single sensory pathway. On page 302 of this issue, Ahissar and colleagues² propose a possible solution, by showing that these different sorts of

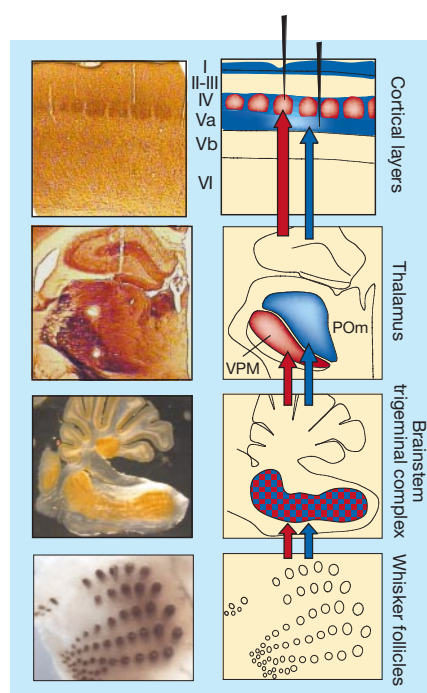


Figure 1 Parallel pathways for touch. Left, histological tissue from various levels of the rat's whisker sensory system. Right, the lemniscal (red) and paralemniscal (blue) sensory pathways, from receptors (bottom) to cortex (top). Bottom, the skin of the snout, where the follicle of each whisker is visible. Sensory fibres lead from here to the brainstem trigeminal complex, where lemniscal and paralemniscal pathways are not yet distinct. Fibres then lead to the thalamus, where neurons in the ventral posterior medial nucleus (VPM) exhibit properties specific to the lemniscal pathway, and neurons in the posterior medial nucleus (POM) show properties specific to the paralemniscal pathway. The two pathways then lead to different cortical layers. Ahissar *et al.*² recorded the activity of neurons in all these brain regions (black lines show the positions of recording electrodes in the cortex). Their results show that the lemniscal pathway may convey information about what is touched by the whiskers. The paralemniscal pathway may encode information about the location of the touched object. Photographs supplied by S. Haidarliu and E. Ahissar.

information are segregated into two sensory channels.

Neuroscience students learn that, in touch, vision and hearing, signals evoked by external events are quickly and reliably channelled to their respective sensory areas in the cortex of the brain through 'relay' neurons in the thalamus (Fig. 1). These channels have been termed 'lemniscal' pathways. But neuroanatomists have long recognized that the lemniscal groups of cells ('nuclei') occupy less than half the volume of the thalamus associated with each type of sensation. The remainder belongs to 'paralemniscal' sensory pathways. The physiological properties of paralemniscal neurons in the thalamus are difficult to classify using typical sensory stimulation protocols. So it has been difficult to assign any specific function to these pathways. But now, Ahissar *et al.*² suggest that the paralemniscal pathway associated with touch information picked up by a rat's whiskers can relay to the cortex information about the position of touched objects.

Each full forwards-backwards sweep of a rat's whiskers is called a whisking cycle. For the brain to obtain any useful information about the location of objects touched during this cycle, it needs to reconstruct the trajectory of the whiskers. Conceivably, the sensory cortex could receive data about the cycle directly from the motor systems that generate the whisker movements. But this does not appear to be the case. Motor cortex, although a major source of input to sensory cortex, does not carry a strong signal about whisker position³. By default, then, information about the whisking cycle must arise through a signal transmitted from the whiskers.

To identify the signal that reveals whisker position, Ahissar *et al.*² induced whisker movement in anaesthetized rats by directing air-puffs along the nose. The whisker deflections, in the range of 2–8 Hz, were taken as experimentally controlled whisking cycles.

The flow of information from whiskers to cortex is shown in Fig. 1. Lemniscal signals are transmitted from receptors at the base of the whiskers, via the brainstem trigeminal complex, to the medial portion of the ventral posterior nucleus of the thalamus (VPM), ending up in the tactile region of cortex. Paralemniscal signals are transmitted along a similar pathway, but through the medial portion of the posterior nucleus (POM) in the thalamus. Ahissar *et al.* looked at the neuronal responses to whisker movement in each of these brain regions. For each frequency of whisker deflection, the authors aligned the neuronal responses across cycles to yield an average response per cycle. They then measured the response latency (the time between the start of whisker movement and the beginning of neuronal activity) and response magnitude (the number of 'spikes' of neuronal activity per whisker movement).

They found that neurons in the trigemi-

nal complex — where the lemniscal and paralemniscal pathways are not yet distinct — fired with an identical response latency and magnitude regardless of whisking frequency. In both VPM and POM, by contrast, the response magnitude decreased as the frequency increased. But in VPM, response latency was constant regardless of frequency, whereas in POM response latency increased as whisking frequency increased. Neurons at the cortical targets of the two pathways (cortical layer IV for VPM, and layer Va for POM) behaved similarly to their thalamic inputs.

Ahissar *et al.* propose that a feedback circuit between sensory cortex and POM is responsible for the latency shifts with increasing whisking frequency. Cortical neurons in layer V have intrinsic oscillatory mechanisms⁴ and a powerful influence on POM⁵. Ahissar *et al.* suggest that POM neurons activate cortical inhibitory neurons; these then inhibit the cortical oscillators, which in turn drive cortex-to-thalamus feedback. The properties of this loop could account for the changing response latency seen in POM.

There is no direct proof that such feedback loops can decode whisker position. But Ahissar *et al.* have shown that this varying response to whisking frequency endows the paralemniscal system with the ability to process temporal information — such as object location — that varies on timescales as slow as single whisking cycles. By comparing the timing of feedback descending from the cortex with the timing of information ascending from the whiskers, POM neurons could signal where along the whisking sweep an object is located. The lemniscal pathway, by contrast, responds with a fixed latency, so it is better suited to representing object features — such as surface texture — that do not vary along the whisking cycle.

Can these ideas be generalized to other

sensations? What must be represented by the brain is unique to each kind of sensation, so one cannot expect the nature of sensory coding to be identical. The tactile sensory pathways form neural representations of objects that contact receptors in the skin; the visual and auditory systems form representations of objects at a distance.

But the observations made by Ahissar *et al.* raise some interesting ideas about how non-tactile stimuli may be represented. For example, neuroscientists debate whether the transmission of information is based on neuronal firing counts over relatively long time windows ('spike-count' coding) or on variations in neuronal firing within brief time windows ('spike-time' coding). Ahissar *et al.* have shown that it may be efficient for these codes to coexist. POM neurons transmit information about the frequency of whisker movements by shifts in response latency — a clear spike-time code. But the response latency determines the time window available for firing, so these timing shifts lead to spike-count differences, too. So frequency information is present in both the total spike count and the spike timing.

Time will tell whether or not these results can be applied to other sensory systems. But it seems that, at least for touch, the parallel pathways long recognized by neuroanatomists are beginning to acquire functional significance. ■

Mathew E. Diamond is in the Cognitive Neuroscience Sector, International School for Advanced Studies, Via Beirut 2-4, 34014 Trieste, Italy.

e-mail: diamond@sisa.it

1. Carvell, G. E. & Simons, D. J. *J. Neurosci.* **10**, 2638–2648 (1990).

2. Ahissar, E., Sosnik, R. & Haidarliu, S. *Nature* **406**, 302–306 (2000).

3. Fee, M. S., Mitra, P. P. & Kleinfeld, D. *J. Neurophysiol.* **78**, 1144–1149 (1997).

4. Silva, L. R., Amitai, Y. & Connors, B. W. *Science* **251**, 432–435 (1991).

5. Diamond, M. E., Armstrong-James, M., Budway, M. J. & Ebner, F. F. *J. Comp. Neurol.* **319**, 66–84 (1992).

Atmospheric physics

Enlightening water vapour

Brian J. Soden

The extent to which the Earth's climate will warm as a result of increasing carbon dioxide in the atmosphere depends largely on the response of water vapour in the upper levels of the troposphere (roughly 5–10 km above the surface). But our ability to monitor changes in water vapour there is limited by the scarcity of observing systems with sufficient accuracy and longevity to document its global variation and to detect trends.

On page 290 of this issue¹, Price demonstrates a remarkably robust relationship between upper tropospheric water vapour

and global lightning activity. This observation not only supports the idea that atmospheric convection moistens the upper troposphere, a characteristic of climate models that has been the subject of spirited debate², but may also provide a unique tool for monitoring purposes.

The importance of water vapour in regulating climate is without question. It is the dominant greenhouse gas, trapping more of the Earth's heat than any other gaseous constituent of the atmosphere. If water vapour concentrations increase in a warmer world, as is widely believed, the added absorption

will act to further amplify the initial warming. Mathematical models of the Earth's climate suggest that this would provide considerable positive feedback, more than doubling the sensitivity of surface temperature to anthropogenic greenhouse gases. Current models predict that water vapour concentrations in the upper troposphere could increase by as much as 40% over the next half-century³. This amplified moistening in response to the predicted warming not only highlights the significance of water vapour as a feedback mechanism, but also underscores the need for long-term monitoring of upper-tropospheric water vapour in helping to detect climate change and identify its causes.

Some climate researchers have challenged this view. They argue that the treatment of water vapour in climate models is overly simplified and that concentrations in the upper troposphere might actually decrease in a warmer climate. The mechanisms and effects of atmospheric convection and related cloud processes have been central to this debate.

Lindzen² postulated that the increased convective overturning of the atmosphere in response to warmer surface temperatures would bring more dry air down from higher altitudes and decrease the moisture in the upper troposphere. That atmospheric convection locally moistens the environment in which it occurs is indisputable. Observations from weather satellites (Fig. 1) clearly show increased levels of humidity (yellow to red) in regions surrounding the areas of active convection, as can be seen from the upper-level cloud cover (grey). However, on a global scale, the net effect of an increase in convective overturning is less obvious. Rising air in convective towers must be balanced in surrounding regions by sinking air, which, in the absence of other sources of moisture, can lead to extremely dry conditions in the upper troposphere (blue in Fig. 1).

Price¹ uses global lightning activity as a proxy for atmospheric convection. This is reasonable because convection not only determines the vertical transport of moisture, but also influences the electrification processes responsible for generating lightning activity. He shows that variations in lightning are well correlated with the observed fluctuations in global upper-tropospheric moisture on both weekly and seasonal timescales. Some may rightly question the reliability of current systems for observing moisture levels, such as the satellite output that Price has used. But he shows that similar variations in water vapour are obtained using independent measurements from different observing systems, which lends further credibility to the results.

By itself, the strong correlation between lightning activity and upper tropospheric water vapour offers a valuable opportunity for testing model simulations of the response of moisture to changes in atmospheric convection. But Price also points out that global

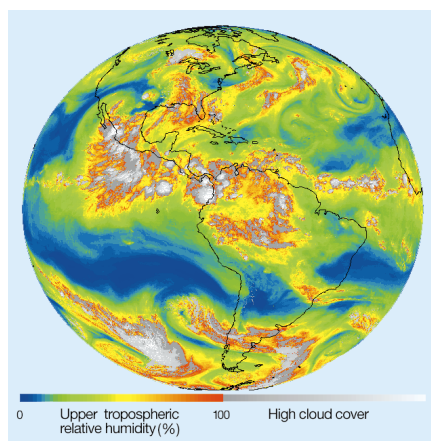


Figure 1 Water vapour in the upper troposphere. This satellite image shows that higher levels of humidity (yellow to red) are associated with the areas of active atmospheric convection evident from clouds (grey). Price¹ takes global lightning as a proxy for convection, and shows that it might be used to monitor water vapour concentrations in the upper troposphere.

lightning activity can be readily monitored from a single location through so-called Schumann resonances, the low-frequency electromagnetic waves discharged by lightning. Because these waves become trapped in the wave guide formed by the Earth's ionosphere, they can travel around the globe several times before dissipating. This means that global monitoring can be done from a single location. Given the close relationship between lightning and upper-tropospheric water vapour, Price suggests that measurements of Schumann resonances could provide an

inexpensive method for monitoring the long-term trends in upper-tropospheric moisture.

Water vapour is notoriously difficult to measure and such a capability would be of great value to the climate research community. Water vapour is highly variable in both space and time (Fig. 1), and its concentration in the atmosphere varies by over three orders of magnitude. Although an international network of weather balloons has carried water vapour sensors for nearly half a century, changes in instrumentation and poor calibration make such sensors unsuitable for detecting trends in upper-tropospheric water vapour. Likewise, although weather satellites have provided global measurements of water vapour for over two decades, little effort has gone into making the data appropriate for long-term climate monitoring.

Price has shown that seasonal changes in global lightning activity are well correlated with changes in water vapour in the upper troposphere, but it remains to be seen whether the long-term trends in these quantities exhibit similar consistency. If Price's optimism proves correct, and future trends in lightning activity can indeed be linked to trends in upper-tropospheric water vapour, observations of Schumann resonances could provide a great deal of enlightenment in tracking climate change.

Brian J. Soden is at the NOAA Geophysical Fluid Dynamics Laboratory, Princeton University, PO Box 308, Princeton, New Jersey 08542, USA.

e-mail: bjs@gfdl.gov

1. Price, C. *Nature* **406**, 290–293 (2000).
2. Lindzen, R. S. *Bull. Am. Meteorol. Soc.* **71**, 288–299 (1990).
3. Rind, D. *Science* **281**, 1152–1153 (1998).

Human behaviour

Fair game

Ruth Mace

Imagine that you are with a group of people from your community, and have been given a sum of money on condition that you share it with an anonymous member of the group. You make them an offer of a proportion of the cash, but neither you nor the other player will get a penny unless they accept your offer. There is only one round of the game, and confidentiality is maintained.

What do you think is a fair offer? The answers to this question, for a range of cultures from around the world, were the subject of a symposium* held last month: economic game theory met human behavioural ecology with fascinating results.

The rational way to play the game, known in economics as the ultimatum game, is for

the first player to offer as little as possible and for the second player to accept it — something is always better than nothing. But when groups of university students play, the most common offer is 50% of the money, with a mean of about 40%. Offers below 20% are almost always rejected. Rejection is pure spite, as no one gains, but it is the only way of punishing the proposer for a mean offer. A range of other experiments confirm a general tendency in humans both to be generous to distantly related individuals and to punish cheaters, sometimes at great personal cost (H. Gintis, Univ. Massachusetts, Amherst).

Western university students are the social scientist's equivalent of the laboratory rat, yet evolutionary psychologists and economists are often tempted to assume that results obtained with them represent universal human preferences. But Westerners, not

* *Reciprocity and Human Sociality*, Human Behaviour and Evolution Society, Amherst College, Massachusetts, USA, 7–11 June 2000.

bound by the constraints under which humans evolved, may not be a good evolutionary model. The first attempt to play the ultimatum game in a traditional society — the remote Machiguenga subsistence farmers of the Amazon (J. Henrich, Univ. California Los Angeles) — highlighted the possibility that there may be cross-cultural variation in the human concept of what is fair. The Machiguenga routinely made and accepted offers of 20% or less: Machiguenga are mean, but still not mean enough to be considered truly rational. Can their relative poverty compared to Westerners explain their behaviour?

Gintis had persuaded several more anthropologists to fan out around the globe and get people from a range of traditional cultures to play the ultimatum game. This often involved logistical headaches: groups of people were sequestered in different parts of a forest to maintain anonymity, with messengers running between them, and players were impounded until all had finished playing so they could not influence each other by discussing their strategies. The results confirmed the early suspicions of cultural differences, in that a bewildering variety of offers were considered fair. But how and why was open to interpretation, and each anthropologist had a different explanation.

The Machiguenga turn out to be at the mean end of the spectrum. A range of other subsistence farmers, as well as hunter-gatherers, living in circumstances where wealth can determine infant mortality and fertility, may both offer and reject large sums of money that are equivalent to weeks of income in some cases. Oromo pastoralists in northern Kenya made generous offers, although those who were less involved in market trading, and thus had less cash, were less likely to do so (J. Ensminger, Washington Univ., St Louis). Hadza hunter-gatherers from Tanzania (F. Marlowe, Harvard Univ.) living in large villages make more generous offers than those from smaller camps. This is counter to the predictions of both kin selection (helping your relatives) and reciprocal altruism (helping individuals who may help you on future occasions) that altruism is likely to be stronger in smaller groups.

Possibly the tendency to share, and also to punish those who do not offer enough, arises from the level of inter-reliance in the group. This mutualism could explain why, in Lamalera, Indonesia, traditional whale-hunters would accept nothing but generous offers, which were usually what were made. Team effort is essential to hunt whales in open rowing-boats (M. Alvard, State Univ. New York, Buffalo) — these hunters are literally all in the same boat.

In Conambo, in the Ecuadorian Amazon, levels of sociality varied within one village that could be divided into two alliances (J. Patton, Washington State Univ., Pullman). Although individuals did not know

which alliance their partner belonged to, those from the more stable alliance were more likely to make generous offers, suggesting that time discounting was a consideration — if you are not likely to stay in a village for long, why be generous?

A common thread to these results is that interactions within human communities may not be played out simply through the assessment of individual costs and benefits (although that surely operates too). Rather, behaviour may be influenced by a general sense of the level of group cohesion, stability and inter-dependence. The tendency to co-operate with members of the group and to punish those who do not cooperate, even at some personal cost, has been described as 'strong reciprocity' (H. Gintis *J. Theor. Biol.*, in the press) — this is a form of altruism that is not easily explained by kin selection,

reciprocal altruism or any of the other evolutionary models of cooperation based on individual advantage.

Such findings have stimulated evolutionary theorists to return to models of the evolution of group-beneficial traits. Several models of the evolution of such behaviour exist. They include those based on cultural (J. Soltis *et al. Curr. Anthropol.* **36**, 473–483; 1995) or possibly genetic traits (Gintis), all in populations structured in small groups at high risk of extinction. Such conditions could have prevailed in human societies in the not-too-distant evolutionary past. But whether this led humans to evolve a genuine social conscience is still open to debate. ■

Ruth Mace is in the Department of Anthropology, University College London, Gower Street, London WC1E 6BT, UK.

e-mail: r.mace@ucl.ac.uk

Structural biology

Adding to the antifreeze agenda

Charles A. Knight

Antifreeze proteins allow certain organisms with no thermal control to survive cold conditions, the best known examples to date coming from fish. Papers elsewhere in this issue^{1–3} now describe the characterization of such proteins from grass and two species of insect. They provide insights into the mysterious ways that antifreeze proteins work.

The biological 'antifreezes' are solutes that stick to ice crystals in supercooled water and prevent their growth. (Supercooling refers to liquids below their freezing points; water is easily supercooled to several degrees Celsius below freezing, and remains liquid indefinitely unless nucleated.) So these molecules are unlike the antifreezes used in internal combustion engines, which lower the freezing point by changing the properties of liquid water. As surface-active materials, small quantities of antifreeze proteins can have large effects on surface phenomena such as crystal growth. Antifreezes from cold-water fish range upwards in relative molecular mass from 2,600, and the most active prevent ice-crystal growth over a supercooling temperature range of just over 1 °C. In combination with dissolved salts, this is enough to protect the fish from freezing in sea water at its freezing point of –1.9 °C.

In the new papers, Sidebottom *et al.*¹ (page 256) report the isolation from grass of a new antifreeze that is an exceptionally good inhibitor of ice recrystallization. Graether *et al.*² and Liou *et al.*³ describe β -helical molecular structures of antifreezes from two insect species; β -helices are found only rarely in

proteins, and have not previously been associated with antifreeze activity.

When I first read about the fish antifreezes 20 years ago, ten years after their discovery, I found it hard to believe that a solute could adsorb at the ice–water interface and completely stop crystal growth. The interface provides no chemical contrast whatever and ice grows in water very readily indeed. Adsorption at interfaces between the molten and crystalline phases of the same material, such as water and ice, had been virtually ignored, and as far as I knew this was the first example of complete adsorption-inhibition of crystal growth from the molten phase. The experiments were quickly confirmed, however, and with that came belief in the principle.

The defining experiment for antifreeze effects on ice-crystal growth was done by A. L. DeVries⁴. Studying the behaviour of tiny ice crystals in solutions of pure antifreezes in capillaries, he found that the ice does not grow observably until a certain supercooling is reached. Then, when the temperature is decreased further, ice needles sprout from the crystal, growing suddenly and swiftly. This 'freezing point' is a reproducible function of solution concentration. On warming, the ice does not melt until the expected melting point is reached, very near 0 °C. The discrepancy between experimental melting and freezing points is known as freezing hysteresis.

Early on, it was thought that antifreezes probably stick to ice by hydrogen bonding aided by a structural fit to the ice surface. The orientation-dependence of adsorption was shown later by an etching technique⁵. (The crystallographic specifics are that antifreezes

from different fish species were found to adsorb at prism, secondary prism and pyramidal orientations, but not on the basal plane. This explains why the ice grows as needles parallel to the *c*-axis of the unit cell.)

Applications of antifreeze proteins, some based on their being potent inhibitors of recrystallization⁶, have been sought for maintaining texture in frozen foods, improving storage of blood, tissues and organs, cryosurgery, and protecting crops from freezing. But just what is it that gives these molecules their activity? One side of them is relatively hydrophobic and the other relatively hydrophilic, and there has been disagreement over which side binds to ice⁷. Much of the evidence has come from freezing hysteresis measurements on synthesized versions of the natural proteins with amino-acid substitutions introduced into them.

The observations of Sidebottom *et al.*¹ raise the issue of the relationship between the two main effects of antifreeze proteins: freezing hysteresis (protection from freezing) and recrystallization inhibition (in organisms, presumably protection from freezing damage once freezing does occur). The general understanding has been that, like freezing hysteresis, recrystallization inhibition is a direct consequence of immobilization of solid-liquid interfaces in partially frozen samples. Sidebottom and colleagues' results suggest, however, that the two effects are uncoupled; if they are, then that understanding is probably flawed.

The molecular structures of antifreeze peptides described by Graether *et al.*² and Liou *et al.*³ are the first from insects to be completely characterized. Both show beautiful structural matches between ice and the hydrophilic groups on one side of the molecule. Most of the fish antifreeze proteins shown to fit onto ice were linear molecules. Here, however, the arrays are two-dimensional in both cases, so there can now be little doubt that it is the hydrophilic sides that stick to the ice, leaving the relatively hydrophobic sides in contact with liquid water.

The structures^{2,3} are similar in being β -helices, showing a new way in which peptide-chain folding can bring about the structural match that seems to be necessary for antifreeze activity. No doubt more ways will be found in the future. The very active spruce-budworm antifreeze of Graether *et al.*² is novel in yet another respect, in that the ice grows as basal plates from solution rather than as needles; this may be the first case of an antifreeze protein binding to the basal face of ice.

What of more general issues? The usual approach to antifreeze adsorption to ice has been to analyse the bonding between the antifreeze molecules and ice. In a different viewpoint, which I favour, the rather large antifreeze molecules can be viewed instead as very small particles. Their equilibrium positions at the ice-water interface can then be considered

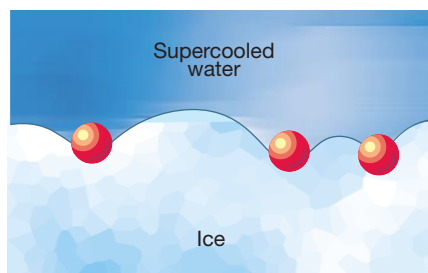


Figure 1 Two-dimensional model of ice-growth inhibition by antifreeze molecules (red). One side of each molecule sticks to the ice, which grows between the molecules. But the curvature lowers the local freezing point through the Kelvin effect.

in terms of interfacial energies, and insights from theoretical studies of the interactions of particles with growing crystals can be applied^{8,9}. The important parameter is the difference in energy between the antifreeze-ice and antifreeze-water interfaces, and the principle is minimization of the total interfacial energy. Accordingly, even if an antifreeze molecule has no preference between contacting ice or water (its interfacial energies with them are the same), it may adsorb strongly because the area of the ice-water interface is reduced when it resides there.

Another point is that antifreeze proteins must match the ice structure well enough to prevent water molecules diffusing into the interface and pushing the protein ahead as the crystal grows. This 'pushing ahead' is the crystallization pressure familiar from many frost-heaving studies¹⁰, in which the particle can maintain its equilibrium position at the surface as the crystal grows. Water molecules diffuse behind the particles, pushing them ahead and forming ice lenses in freezing soil. Considerable pressures are generated in this way, even over a supercooling temperature range of 1 °C or less.

With an array of antifreeze molecules anchored permanently to the ice and unable to be pushed ahead, ice growth only occurs at a supercooling sufficient to engulf the molecules, which then depends on their spacing on the ice surface. The depressed 'freezing point' presumably comes from the Kelvin effect, the local decrease in freezing point at the curved ice interface that bulges forward between the adsorbed antifreeze molecules (Fig. 1). The three-dimensional geometry of this is complicated, however, and has not been worked out.

The question that seems hardest to answer is that, if the adsorption is effectively permanent (as is necessary to stop the ice growth), why does the observed freezing point depend on the solution concentration? Especially at low solution concentrations, one would think it ought to depend more on the time available for antifreeze molecules to adsorb, but experimentally it does not seem to.

All in all, the three papers¹⁻³ discussed here do indeed provide thought-provoking information on the structure and behaviour of antifreeze proteins. But clearly there remains a lack of consensus on some of the fundamental issues to do with antifreeze action.

Charles A. Knight is at the National Center for Atmospheric Research, PO Box 3000, Boulder, Colorado 80307-3000, USA.

e-mail: knightc@ncar.ucar.edu

1. Sidebottom, C. *et al.* *Nature* **406**, 256 (2000).
2. Graether, S. P. *et al.* *Nature* **406**, 325-328 (2000).
3. Liou, Y.-C., Tocilj, A., Davies, P. L. & Jia, Z. *Nature* **406**, 322-324 (2000).
4. DeVries, A. L. *Science* **172**, 1153-1155 (1971).
5. Knight, C. A., Cheng, C. C. & DeVries, A. L. *Biophys. J.* **59**, 409-418 (1991).
6. Knight, C. A., DeVries, A. L. & Oolman, L. D. *Nature* **308**, 295-296 (1984).
7. Haymet, A. D. J., Ward, L. G. & Harding, M. M. *J. Am. Chem. Soc.* **121**, 941-948 (1999).
8. Uhlmann, D. R., Chalmers, B. & Jackson, K. E. *J. Appl. Phys.* **35**, 2986-2993 (1964).
9. Asthana, R. & Tewari, S. N. *J. Mater. Sci.* **28**, 5414-5425 (1993).
10. Jackson, K. A. & Chalmers, B. *J. Appl. Phys.* **29**, 1178-1181 (1958).

Mathematics

Best packing in proteins and DNA

Andrzej Stasiak and John H. Maddocks

A pragmatic way to store a rope is to coil it loosely, and drop it in an appropriately shaped box. But if you are extremely parsimonious with space, as nature often is, this solution is suboptimal — there is a lot of unused space in the centre of the coil. So what is the longest piece of rope you can pack into a particular box? Such questions of optimal packing are addressed by Maritan *et al.*¹ on page 287 of this issue. Some of the optimal shapes they find are the familiar, naturally occurring, helical structures of proteins and DNA.

Optimal packing is a classic area of dis-

crete geometry. Perhaps the best known example is Kepler's problem of the densest packing of identical spheres², which has important applications in such fundamental physical phenomena as crystallization and melting of condensed matter. But most physical objects cannot be modelled as simple spheres. In particular, many polymers, including DNA molecules and portions of proteins, might more reasonably be modelled as deformable tubes. For example, if we consider how DNA molecules could be packed within a small virus³, we arrive at the question of optimal packing of tubes. Unlike



100 YEARS AGO

It is a remarkable sign of the times when the head of a firm principally distinguished for the introduction into this country of American methods of dealing with drugs, i.e. by putting them up in new and convenient shapes and doses, goes out of his way to fit up extensive research laboratories. This is what Mr. Wellcome has done... A well-built modern house has been secured at No. 6 King Street, Snow Hill, and has been converted into a series of three commodious and well-fitted laboratories, a library and office, and a store-room and workshop. Each laboratory is self-contained and each is connected with the other and with the directors' office by means of telephones... Mr. Wellcome intends to carry on his laboratories in no narrow spirit; this means, I presume, that he has other views than the conversion of his business into a chemical manufacturing concern. Though much work is done towards the perfection of the firm's preparations, time has been found for several researches which have been published, and other work of this kind is in hand... All interested in the advance of chemistry, whether pure or applied, will wish Mr. Wellcome success, and also that he may find imitators among the numbers of firms who are meditating an advance in the direction of a more scientific method of conducting their manufactures. From *Nature* 19 July 1900.

50 YEARS AGO

Crystalline inclusion bodies in tobacco plants infected by tobacco mosaic virus have been known since 1903, and circumstantial evidence has made it appear likely that these crystals are composed largely of the virus protein. The present work makes it appear even more likely than before that the crystals are pure virus protein, and shows the crystals to be of considerable interest from several quite different but related points of view. On account of the exceptionally large dimension of the protein particle, it has been possible for the first time to make, in part at least, a structure analysis of the crystal using visible light in a manner analogous to that of X-ray diffraction. As a result, it has been possible to settle the controversial question of the length of the rod-shaped virus particle in the living plant. Also, the interpretation of the appearance of the crystals, as seen with the microscope, leads to a theory of the formation of images of three-dimensional objects. M. H. F. Wilkins *et al.* From *Nature* 22 July 1950.

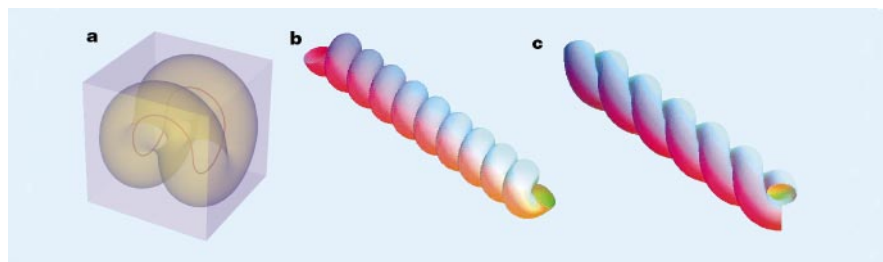


Figure 1 Tightly packed tubes. a, Packing of a tube into a box whose sides are four times the radius of the tube. b, A single helical tube whose centre line has the critical pitch/radius ratio of 2.512 . c, Double-helical tubes in which the ratio between pitch and tube radius has the critical value 2π . Helices with parameters as in b closely resemble those observed in α -helical segments of proteins and are also obtained in numerical simulations of optimal packing of individual tubes by Maritan *et al.*¹, whereas the parameters of the double-helical tubes in c closely match those for DNA.

spheres whose geometry is fixed, the optimal shape of the centre line of a deformable tube must be found. Consequently it is not a simple matter to find the optimal packing of even a single tube.

This modelling problem makes sense only if the tube, like any real rope or molecule, has a definite non-zero thickness. Figure 1a shows a way of packing (presumably optimally) a tube of length 4π and uniform thickness r (or radius r , but be aware that sometimes thickness means diameter) into a cubic box of side length $4r$. The centre line of the tube is a saddle-like trajectory that is built from four semi-circles. The point of this example is that, whenever the parameters of the tube and the box are less simply related, it is hard to even guess the optimal packing. This is the general topic addressed by Maritan *et al.*¹ in their numerical simulations.

The basic issue to be overcome in numerical simulations is that the idea of thickness is not quite as simple as it might first seem^{4,5}. In the tube model a non-zero thickness has two possible effects: first, the centre line cannot be bent too sharply; and second, any two points that are far apart along the curve cannot be too close to each other in space. So, for a given centre line, the maximum possible thickness is governed by either local bending or by non-local points of closest approach, or, apparently exceptionally, by both conditions simultaneously. For the centre line of the tube shown in Fig. 1a, both of these conditions are simultaneously realized at every point along the red centre line.

Maritan *et al.* characterize the thickness of strings or tubes in terms of a quantity called global radius of curvature⁶. For any curve made up of discrete straight lines joining node points, the thickness is taken to be the minimal radius of all possible circles passing through any three nodes of the curve. If the smallest radius is achieved by three adjacent nodes, the thickness is controlled by local bending, whereas if the nodes on the minimal circle are not all adjacent, the thickness is governed by the non-local condition of closest approach.

Armed with this tool, Maritan *et al.* use a

Monte Carlo algorithm to move the nodes of a centre line with a given length into an optimal shape, which maximizes thickness when subject to one of several compactness constraints. Perhaps the simplest compactness condition is that the centre line is completely contained in a given box. In spirit, their procedure is similar to that of earlier studies (see examples in ref. 7) but, with the exception of ref. 8, all previous work has looked at the optimal shapes of closed, knotted curves, much beloved of mathematicians. (A curve is closed if its two ends are joined or glued together to form a loop. A loop is knotted if it cannot be smoothly deformed to a simple circle without cutting.) Indeed, from a solely mathematical point of view, Maritan *et al.*'s contribution is the elegant, and in retrospect delightfully obvious, idea that the constraints of closure and knotting can usefully be replaced by one of several compactness conditions on the centre line. (In the absence of any compactness constraint at all, the optimal centre line is merely a straight line of infinite thickness.)

What about the physical implications of the simulations presented by Maritan *et al.*? Perhaps their most intriguing results arise when they impose local compactness conditions that are independent of external constraints such as a box. They state that, for "a broad class of local constraints", the optimal centre line is a particular helix in which the ratio of the pitch (or period) p to radius r is such that the bending and closest-approach constraints are realized everywhere simultaneously ($p/r = 2.512$). Maritan *et al.* then consider crystal structures of various α -helical polypeptides (one of the basic structural motifs of proteins), and show that the helices formed by the α -carbons in the polypeptide backbone have almost the same optimal shape as found in their simulations (Fig. 1b).

Are optimal packings of tubes related to other basic structural motifs in biology? For example, does the DNA double helix also involve optimally packed tubes? A related problem studied by Pieranski⁸ involves finding the densest coiling of two identical

interwound tubes forming a double-helical structure (Fig. 1c). As noted by Pieranski, for this problem there is a critical pitch-to-radius ratio ($p/r = 2\pi = 6.28$) above which the line of contact between the tubes is straight, and below which it is helical. The crystallographic diameter of the classic DNA double helix is 23.7 Å, which would correspond to a radius for each of the two idealized helical tubes of $23.7/4 = 5.92$ Å (here the sugar-phosphate backbones lie on the outside of the helical tubes). The generally accepted value for the pitch is 10.5 base pairs or 35.7 Å, so $p/r = 35.7/5.92 = 6.03$, which is within 4% of 2π . Is this an insight or just a coincidence? We suspect it is the former.

Andrzej Stasiak is at the Laboratoire d'Analyse

Ultrastructurale, Université de Lausanne, CH-1015 Lausanne, Switzerland.

e-mail: andrzej.stasiak@lau.unil.ch

John H. Maddocks is in the Department of Mathematics, Swiss Federal Institute of Technology Lausanne, CH-1015 Lausanne, Switzerland.

e-mail: maddocks@dma.epfl.ch

1. Maritan, A., Micheletti, C., Trovato, A. & Banavar, R. B. *Nature* **406**, 287–290 (2000).
2. Sloane, N. J. *Nature* **395**, 435–436 (1998).
3. Odijk, T. *Biophys. J.* **75**, 1223–1227 (1998).
4. Katritch, V. *et al.* *Nature* **384**, 142–145 (1996).
5. Litherland, R., Simon, J., Durumeric, O. & Rawdon, E. *Topol. Appl.* **91**, 233–244 (1999).
6. Gonzalez, O. & Maddocks, J. H. *Proc. Natl Acad. Sci. USA* **96**, 4769–4773 (1999).
7. *Ideal Knots* (eds Stasiak, A., Katritch, V. & Kauffman, L. H.) (World Scientific, Singapore, 1998).
8. Pieranski, P. in *Ideal Knots* (eds Stasiak, A., Katritch, V. & Kauffman, L. H.) 20–41 (World Scientific, Singapore, 1998).

Cell biology

A protein accumulator

Jennifer C. Pinder and Anthony J. Baines

We thought we knew what spectrin does. Is it not the elastic, membrane-bound protein that prevents red blood cells from rupturing as they circulate in the bloodstream? And does it not have the same supporting function in other cells? The second assumption has seldom been questioned over the past two decades, but has just been overturned by the power of experimental genetics, as described in three reports^{1–3} in the *Journal of Cell Biology*. The results may bear on human diseases such as muscular dystrophy.

Red blood cells with deleterious mutations or deficiencies in spectrin have weakened outer membranes, are misshapen and lack resilience⁴, so it is not surprising that spectrin has long been assumed to be necessary for membrane integrity. Over the years, other functions have been ascribed to spectrin as well. For example, it is thought to be

involved in generating the polarized morphology of epithelial cells, and to have roles in the functioning of the Golgi complex — an organelle involved in protein secretion from cells — and the organization of synaptic vesicles (see, for example, refs 5–7).

To probe spectrin's function, Hammarlund *et al.*¹ and Moorthy *et al.*² have taken advantage of the simplicity of the genome of the nematode worm *Caenorhabditis elegans*. This genome, like that of the fruitfly *Drosophila melanogaster*, has only three spectrin genes, encoding α , β and β_H forms of the protein (Fig. 1). Hammarlund *et al.*¹ looked at worms with a mutation called *unc-70*, which they discovered to lie in the β -spectrin gene. Moorthy *et al.*² used a now common technique for blocking protein expression: they injected double-stranded RNA into the worm's gonad to block the expression of one or more spectrin subunits

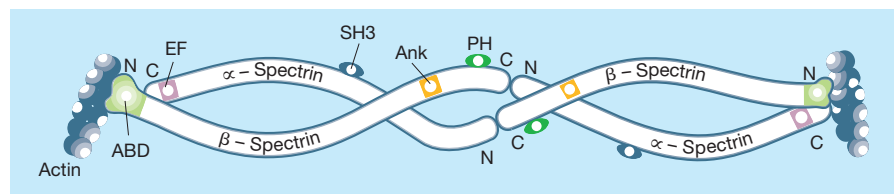


Figure 1 The structure of spectrin^{12,13}. Spectrin is a giant molecule comprising α and β subunits, of which there are different types. For example, *Drosophila* and *C. elegans* have α , β and β_H forms of spectrin. (β_H -Spectrin is a 'heavy' form of β -spectrin.) The α and β subunits associate to form an elongated $(\alpha\beta)_2$ tetramer. Lying near to the interior surface of the plasma membrane, spectrin forms a hexagonal lattice, the nodes of which are crosslinked by the cytoskeletal protein actin. This network is attached to the membrane in several ways, for example through the connecting protein ankyrin. β -Spectrins have binding sites for α -spectrin, actin and ankyrin (Ank). The pleckstrin-homology (PH) domain binds to certain membrane lipids. The Src-homology-3 (SH3) domain of α -spectrin probably accumulates signalling molecules close to the membrane. EF-hands are calcium-binding sites. ABD, actin-binding domain; C, carboxy terminus; N, amino terminus.

individually and in combination. They then analysed the resulting embryos.

Surprisingly, both groups find that β -spectrin is not essential for many of the processes suggested from previous investigations. Their worms do not lose general membrane integrity, and synaptic vesicles in nerve endings are clustered normally. The cellular secretory pathways appear unimpaired: the worms deposit cuticle and secrete collagen and components of the basement membrane as usual. The cells that should be polarized are polarized, suggesting that β -spectrin has no primary function in this process. Ankyrin is a connecting protein that is known to link spectrin to a variety of transmembrane proteins, including cell-adhesion molecules of the L1 family. Moorthy *et al.*² find that, in their worms, ankyrin apparently binds to L1 adhesion molecules normally.

However, the worms are paralysed and have a 'dumpy' appearance¹. The main defects lie in the organization of muscle and nerve cells. The number of neurons is normal, but the patterns of axon outgrowth are altered^{1,2}, with few axons finding their targets. The muscle cells have disrupted sarcomeres (contractile units)^{1,2}, and the sarcoplasmic reticulum — an intracellular calcium store — is generally missing. It seems that the dumpy appearance is caused by a failure of the muscles to spring back after contracting. The muscle defects may arise during the course of contraction, as worms with both the *unc-70* mutation and the *unc-54* mutation, which results in a failure in muscle contraction, show less severe characteristics than the *unc-70* mutants¹.

What, then, has become of the anticipated functions of spectrin? Dubreuil *et al.*³ have looked at fruitflies that lack β -spectrin, and provide some clues to the role of this protein in cell polarization. These flies live just long enough for 'copper' cells in the intestine to be analysed. In the mutant flies, the copper cells are polarized and ankyrin associates with the membrane of these cells as normal. So spectrin is not the main driver of cell polarization. It has been suggested⁸ that what drives cell polarization is the contact of transmembrane cell-adhesion molecules with either their extracellular matrix ligands or their counterparts on other cells. So, in this case, a transmembrane L1-type adhesion molecule binds to its ligand and recruits ankyrin. But a transmembrane ion pump, the Na^+/K^+ ATPase, does not accumulate as normal in the plasma membrane of the mutant copper cells. Why is this?

Spectrin has been described as a 'protein-sorting machine'⁹. The new data^{1–3} show that this model can be taken a step further: spectrin not only sorts, but also collects, proteins at the plasma membrane. Genetic evidence indicates that spectrin functions as a tetramer (Fig. 1). Each tetramer has two

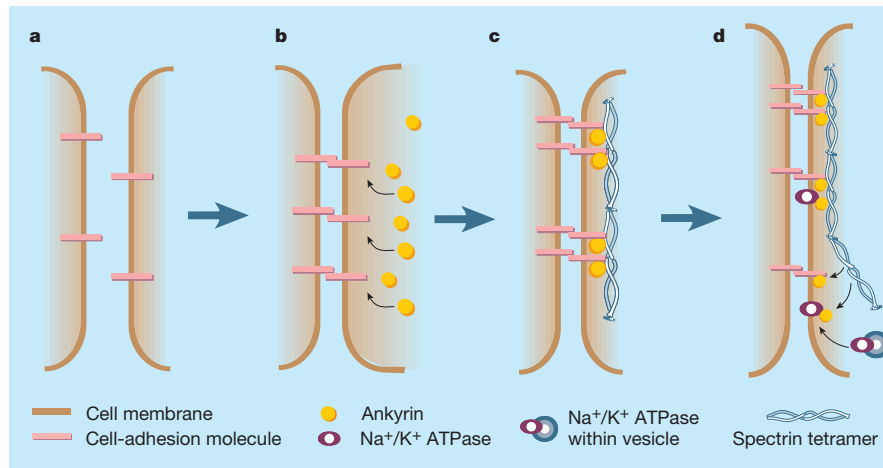


Figure 2 Model for the accumulation of membrane proteins at points on the cell surface specified by cell-adhesion molecules. The new work¹⁻³ is consistent with the following model. **a**, Two cells interact through adhesion molecules. **b**, Ankyrin is recruited to ligand-bound cell-adhesion molecules. Other ankyrin-binding molecules, such as Na^+/K^+ ATPase, do not accumulate in the absence of spectrin. **c**, Spectrin tetramers are recruited to membrane-bound ankyrin/cell-adhesion-molecule complexes, crosslinking these complexes and stabilizing the regions of cell-cell contact. **d**, Further spectrin molecules are recruited into a submembrane skeleton crosslinked by the cytoskeletal protein actin. Binding sites on spectrin trap ankyrin bound to other molecules such as Na^+/K^+ ATPase, and promote their stable incorporation into the membrane domain.

ankyrin-binding sites. One ankyrin molecule, recruited to a cell-adhesion molecule, binds to a spectrin tetramer (Fig. 2). This tetramer then has another ankyrin-binding site free to connect to other ankyrin-containing complexes, such as that consisting of ankyrin and Na^+/K^+ ATPase. If spectrin is missing, this ion pump cannot be captured at the membrane. So spectrin traps and stabilizes proteins at specific points on cell surfaces first specified by ligand-bound cell-adhesion molecules (Fig. 2). The results of Dubreuil *et al.*³ support this theory as far as the Na^+/K^+ ATPase goes. And the results of Hammarlund *et al.*¹ and Moorthy *et al.*² indicate that spectrin may also trap and stabilize key muscle or nerve proteins at specific points on cell membranes as specified by cell-adhesion molecules.

There are more human genes encoding spectrins than there are worm or fruitfly spectrin genes. So, different spectrin isoforms may have evolved other tasks in vertebrates — for example, in Golgi function. This might account for some of the discrepancy between past and present results. But the new results¹⁻³ may also bear on human biology. It is interesting that, in *C. elegans*, a mutation that causes a loss of muscle contraction suppresses the problems that arise from a lack of β -spectrin. This result hints at the idea that any muscle that is in continual use (for example, the heart or diaphragm) might be the most vulnerable to the effects of spectrin mutations, and some unexplained human muscle and heart disorders may have their origins in mutated spectrin genes.

The new work might apply even more generally. Dystrophin is a protein that is mutated in Duchenne and Becker muscular

dystrophies in humans and is a member of the protein superfamily to which spectrin belongs. *C. elegans* with the *dys-1* mutation lack a functional form of dystrophin, but do not show any obvious muscular degeneration. Instead, they lose the enzyme acetylcholinesterase, presumably from its anchoring points in muscle-cell membranes¹⁰. Dystrophic *mdx* mice, which also lack functional dystrophin, lose nitric oxide synthase from the plasma membrane of muscle cells¹¹. Perhaps dystrophin and other members of this superfamily emulate spectrin's protein-accumulating activity.

Jennifer C. Pinder is at the MRC Muscle and Cell Motility Unit, King's College London, New Hunt's House, Guy's Campus, London SE1 1UL, UK.

e-mail: jennifer.fordham@kcl.ac.uk

Anthony J. Baines is in the Department of Biosciences, University of Kent, Canterbury CT2 7NJ, UK.

e-mail: a.j.baines@uk.ac.uk

1. Hammarlund, M., Davis, W. S. & Jorgensen, E. M. *J. Cell Biol.* **149**, 931–942 (2000).
2. Moorthy, S., Chen, L. & Bennett, V. *J. Cell Biol.* **149**, 915–930 (2000).
3. Dubreuil, R. R., Wang, P., Dahl, S., Lee, J. & Goldstein, L. S. *J. Cell Biol.* **149**, 647–656 (2000).
4. Tse, W. T. & Lux, S. E. *Br. J. Haematol.* **104**, 2–13 (1999).
5. Sikorski, A. F., Sangerman, J., Goodman, S. R. & Critz, S. D. *Brain Res.* **852**, 161–166 (2000).
6. De Matteis, M. A. & Morrow, J. S. *J. Cell Sci.* **113**, 2331–2343 (2000).
7. Hofer, D., Jons, T., Kraemer, J. & Drenckhahn, D. *Ann. NY Acad. Sci.* **859**, 75–84 (1998).
8. Yeaman, C., Grindstaff, K. K. & Nelson, W. J. *Physiol. Rev.* **79**, 73–98 (1999).
9. Beck, K. A. & Nelson, W. J. *Am. J. Physiol.* **270**, C1263–C1270 (1996).
10. Giuglia, J., Gieseler, K., Arpagaus, M. & Segalat, L. *FEBS Lett.* **463**, 270–272 (1999).
11. Brenman, J. E. *et al.* *Cell* **82**, 743–752 (1995).
12. Viel, A. & Branton, D. *Curr. Opin. Cell Biol.* **8**, 49–55 (1996).
13. Banuelos, S., Saraste, M. & Carugo, K. D. *Structure* **6**, 1419–1431 (1998).

Daedalus

Pulses of the mind

The human brain, which works by processing discrete nerve impulses, is often regarded as a digital computer. Daedalus wants to measure its processing power in operations per second. Modern magnetic-resonance imaging can detect a region of high brain activity by its increased metabolic rate, but cannot deduce the corresponding digit rate. Daedalus now has an improvement.

He points out that the basic processing unit of the whole nervous system is the nerve pulse. At each point it involves a radial movement of sodium ions lasting about a millisecond. Sodium nuclei have spin and magnetic moment. In a suitable magnetic field, a resonating radio-frequency can elevate them to an excited state, from which they will later relax back to the ground state. Suppose that radio-frequency is 1 kilohertz, corresponding to a time constant of a millisecond. Then relaxation should be decidedly stimulated if the nuclei are being vibrated with the same time constant — as by being in a nerve along which a pulse is passing.

The magnetic field for which sodium ions resonate at 1 kHz is a modest 0.9 microtesla. Daedalus's nuclear magnetic psychometer places the subject's head in such a field and irradiates it at 1 kHz. It measures the relaxation times of the sodium ions in the various regions of his brain. The greater his mental activity, the faster the ions will relax; or more exactly, the stronger the peak of their relaxation-time spectrum within the 1 ms region. Once properly calibrated, the instrument will give the total rate of working for the subject's brain, in pulses per second.

Psychology will at last have a sound numerical basis. Daedalus expects that subjects with a high IQ will show a greater processing rate than those of low IQ. But asked to solve a problem, their rate of working will rise less — their algorithms will be more efficient. Yet intuitive types with an unexceptional IQ may still show high firing rates, from their active imaginations. As the subject learns a new skill, the firing rate in the relevant brain region will rise, and then decline as he automates his new ability. Skilled meditators may be able to drive their brain activity right down; but even a sleeping or drugged subject will still show the processing needed for the brain's steady 'housekeeping' — maintaining heart rate, temperature, peristalsis and so on. All the values should exceed a billion pulses per second, showing up modern computers for the primitive devices they are. **David Jones**

Spider manipulation by a wasp larva

A parasitic wasp forces its host to weave a special web for its own ends.

On the evening that it will kill its orb-weaving spider host, the larva of the ichneumonid wasp *Hymenoepimecis* sp. induces the spider to build an otherwise unique 'cocoon web' to serve as a durable support for the wasp larva's cocoon. The construction of this cocoon web is highly stereotyped, consisting of many repetitions that are almost identical to the early stages of one subroutine of normal orb weaving, the other components of which are repressed. Here I investigate this activity and show that the mechanism employed by the larva to manipulate the spider's behaviour is fast-acting, apparently chemical, and has long-term effects.

The female *Hymenoepimecis* sp. wasp attacks the spider *Plesiometa argyra* at the hub of its orb, stings it into temporary paralysis and lays an egg on the spider's abdomen¹. Subsequently, the spider resumes normal activity. During the next 7–14 days, it builds apparently normal orbs (Fig. 1a) to capture prey, while the wasp's egg hatches and the larva grows by sucking the spider's haemolymph. On the night that it will kill its host, the larva induces the spider to build a cocoon web, moults, kills and consumes the spider, and then spins its pupal cocoon hanging by a line from the cocoon web.

Apparently undamaged cocoon webs almost always had several lines (mean, 3.8 ± 1.4 ; range, 2–8; $n = 39$) radiating in a plane from a 'hub' (Fig. 1b–f). Most radial lines branched near their tips and were attached to the substrate at many points (Fig. 1b). Most webs had only radii and lines connecting them at the hub, and lacked circular hub lines (86% of 66) or frame lines (65% of 77) connecting the radial lines. Any frame lines present were typically much shorter and nearer the hub than those of normal orbs (Fig. 1e). The central portion of the hub was never empty, as it is in a normal orb (Fig. 1a). The most elaborate cocoon webs, however, had distinct hub loops, frames, and a mesh above and below the hub (Fig. 1g), confirming that they were modified orbs.

The construction of cocoon webs was very consistent, and it appeared to be the same as the early stages of type 'D' frame construction (ref. 2, and W.G.E., manuscript in preparation). Many other integral features of normal orb-web construction, including breaking and then reeling up and replacing lines, and breaking and then re-attaching lines^{2–6}, were completely absent. A single 'mistake' in this respect by the spider could be disastrous for the wasp larva, as



Figure 1 Webs of the orb-weaving spider *Plesiometa argyra*. **a**, Prey-capture orb of a mature female; **b**, cocoon web and wasp cocoon from above; **c**, hub of the cocoon web; **d**, cocoon web and cocoon from the side; **e**, cocoon web with one frame line; **f**, the simplest cocoon web, with only two radial lines (larva rests at the hub, consuming the dead spider); **g**, the most complex cocoon web, with circular lines at the hub and a mesh above and below the radial lines. Scale bars for **a–g** are 4, 2, 1, 2, 2, 2 and 2 cm, respectively.

the many-stranded cable of radial lines would be replaced with a single pair of drag lines.

The temporary and sticky spirals were also missing, and the central portion of the finished hub was not removed^{2–6}. These differences between cocoon webs and normal orbs all make the cocoon web a stronger, more durable support for the wasp's cocoon. The importance of this is illustrated by the vulnerability of pupae of the related wasp *H. robertsae* to heavy rains⁷.

The larva makes small holes in the spider's abdomen to imbibe haemolymph¹. As the spider continues to build the cocoon web even when the larva is removed shortly before construction would normally start, the changes in the spider's behaviour must be induced chemically rather than by direct physical interference. The effects are both

rapid (removal earlier in the evening did not result in the formation of typical cocoon webs) and long-lasting (spiders from which larvae were removed built similar webs the following night, although some slowly reverted to more normal orbs on subsequent nights).

The larva's ability to induce specific behaviour patterns in the spider indicates that, at some level within the host, such fine behavioural details are independent units, and not artificial constructs. This point is crucial in the use of behavioural patterns as taxonomic characters^{8,9}.

Many parasites manipulate their host's behaviour^{10–13}, but most of them, particularly insect parasitoids, induce only simple changes, such as movement from one habitat to another, eating more or less, or sleeping^{12,13}. These changes may be instigated by

relatively straightforward mechanisms, for example by the modification of particular receptors^{14,15}. *Hymenopimectis*'s manipulation of its spider host is probably the most finely directed alteration of behaviour ever attributed to an insect parasitoid.

William G. Eberhard

Smithsonian Tropical Research Institute, and
Escuela de Biología, Universidad de Costa Rica,
Ciudad Universitaria, Costa Rica
e-mail: weberhar@cariari.ucr.ac.cr

1. Eberhard, W. G. *J. Hymen. Res.* (in the press).
2. Eberhard, W. G. *J. Arachnol.* **18**, 205–234 (1990).
3. Tillet, A. *La Toile Géométrique des Araignées* (Presses Univ. France, Paris, 1942).
4. Eberhard, W. G. *Evolution* **36**, 1067–1095 (1982).
5. Coddington, J. A. in *Spiders: Webs, Behavior and Evolution* (ed.

- Shear, W.) 319–363 (Stanford Univ. Press, Palo Alto, CA, 1986).
6. Witt, P. N., Reed, C. & Peakall, D. B. *A Spider's Web* (Springer, New York, 1968).
7. Fincke, O., Higgins, L. & Rojas, E. *J. Arachnol.* **18**, 321–329 (1990).
8. Griswold, C. E., Coddington, J. A., Hormiga, G. & Scharff, N. *Zool. J. Linn. Soc.* **123**, 1–99 (1998).
9. Wenzel, J. W. *Amu. Rev. Ecol. Syst.* **23**, 361–381 (1992).
10. Holmes, J. C. & Bethel, W. M. in *Behavioural Aspects of Parasite Transmission* (eds Channing, E. U. & Wright, C. A.) 123–149 (Academic, London, 1972).
11. Barnard, C. J. & Behnke, J. M. *Parasitism and Host Behaviour* (Taylor & Francis, London, 1990).
12. Toft, C. A., Aeschlimann, A. & Bolis, L. *Parasite-Host Associations* (Oxford Univ. Press, New York, 1991).
13. Godfray, H. C. J. *Parasitoids: Behavioral and Evolutionary Ecology* (Princeton Univ. Press, Princeton, NJ, 1994).
14. Jenni, L., Molyneux, D. H., Livesey, J. L. & Galun, R. *Nature* **283**, 383–385 (1980).
15. Wicker, W. Z. *Tierpsychologie* **42**, 200–214 (1976).

Phytochemistry

Heat-stable antifreeze protein from grass

We have discovered an antifreeze protein¹ in an overwintering perennial ryegrass, *Lolium perenne*. The protein is stable at 100 °C and although it is a less effective antifreeze than proteins found in antarctic fish and insects, it is better at preventing ice recrystallization. This property enables grasses to tolerate ice formation in their tissues without being damaged, suggesting that the control of ice-crystal growth rather than the prevention of freezing may have evolved to be the critical factor in their survival at very low temperatures.

Frost-tolerant plants undergo a process of cold acclimation^{2,3}, during which perennial grasses accumulate a boiling-tolerant protein that inhibits ice recrystallization. We extracted the protein responsible for this activity from cold acclimated leaves of *L. perenne* and cloned its complementary DNA by using the polymerase chain reaction⁴. We found that it had an open reading frame encoding a protein of 118 amino acids (GenBank accession number, AJ277399) and relative molecular mass 11.765K, with six potential N-glycosylation sites containing the conserved N-X-S/T glycosylation motif.

Although this boiling-tolerant antifreeze protein (AFP) belongs to a new class of plant proteins and shares no lengthy sequence homology with any other AFP or protein sequence, some of its properties fit with the general pattern for AFPs. It is very hydrophilic, being rich in asparagine (25%), valine (16%), serine (15%) and threonine (10%), and having very few amino acids with aromatic or hydrophobic side chains. The primary structure has a series of highly conserved, 7-amino-acid repeat sequences with regularly spaced serine and threonine residues that may be

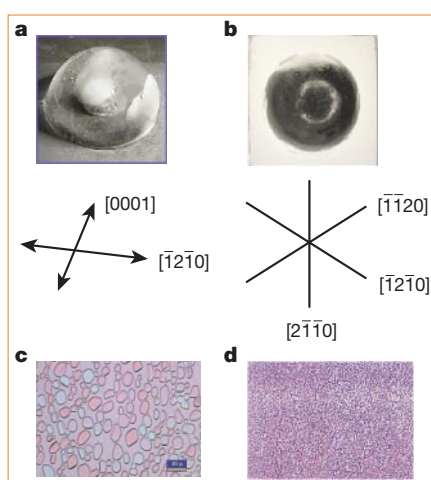


Figure 1 *Lolium* antifreeze protein (AFP) binding to ice and its effect on ice recrystallization. *Lolium* AFP binds specifically to an ice-crystal surface with six-fold symmetry. **a,b**, Ice-etching determination of the binding planes using the hemisphere technique⁵: **a**, three elongated patches positioned on the primary prism plane; and **b**, the planes symmetrically arranged around the crystal's c-axis. **c,d**, Influence of *Lolium* AFP on ice recrystallization. The recrystallization inhibition assay shows crystal growth after 60 min at -6 °C; *Lolium perenne* (**d**) inhibits recrystallization of ice in dilute concentrations relative to growth with the 30% sucrose control (**c**). Scale bar, 50 μm.

able to hydrogen-bond with an ice surface. Growth of a single ice-crystal hemisphere from a dilute solution of the protein, and subsequent surface-etching of the ice hemisphere⁵, produced a distinctive pattern with six-fold symmetry, demonstrating that the protein was specifically binding to ice on the primary prism plane (Fig. 1a,b).

The Fourier-transform infrared spectrum of this grass AFP in solution at room temperature revealed a high solvent-exposed β-sheet content which may be exposed at the ice-binding surface⁶, as proposed for several other antifreeze proteins, including that from carrot and types II and III from fish. The spectrum was the same in the presence of ice, suggesting that the conformation of the *Lolium* AFP does not change on binding to ice, unlike that of the

insect *Dendroides canadensis* thermal-hysteresis protein, which does⁷.

Our *Lolium* antifreeze protein had a significantly higher specific activity in an ice-recrystallization inhibition assay⁸ than other antifreeze proteins. Growth of ice crystals in 30% sucrose solution was completely inhibited at AFP concentrations below 10 μg ml⁻¹ (Fig. 1c,d), which is at least 200 times less in molar terms than the type III AFP from ocean pout (*Macrozoarces americanus*).

In contrast, *Lolium* AFP shows a low thermal hysteresis (the lowering of the temperature at which ice forms on cooling while the melting temperature remains unaltered⁹), with the highest measurable value being 0.1 °C in water and 0.45 °C in 30% sucrose; these values are much lower than the 1.0–1.5 °C reported for fish AFP¹⁰ and the 5–6 °C reported for insect proteins¹¹, although these too are increased by sucrose¹².

Mechanisms previously proposed to explain how AFPs work all imply some correlation between thermal hysteresis effects and recrystallization inhibition¹³. Our discovery that there is no correlation between these relative activities of the antifreeze protein from *L. perenne* raises questions about the nature of the AFP mechanism and indicates that different classes of AFP may interact with ice in different ways. We propose that the thermal hysteresis activity of the grass protein is unlikely to serve an important protective function at the very low temperatures survived by overwintering grasses, whereas its capacity to control the growth of ice crystals may protect it against damage to the plant cellular structure.

**Chris Sidebottom, Sarah Buckley,
Paul Pudney, Sarah Twigg, Carl Jarman,
Chris Holt, Julia Telford,
Andrew McArthur, Dawn Worrall*,
Rod Hubbard*, Peter Lillford**

Unilever Research, Colworth House, Sharnbrook,
Bedford MK44 1LQ, UK

e-mail: Chris.Sidebottom@unilever.com

*Plant Laboratory, Biology Department,
University of York, York YO1 5YW, UK

1. Feeney, R. E. & Yeh, Y. *Food Technol.* **47**, 82–88 (1993).
2. Thomashow, M. F. *Adv. Genet.* **28**, 99–131 (1990).
3. Guy, C. L. *Amu. Rev. Plant Physiol.* **41**, 187–223 (1990).
4. Frohman, M. A., Dush, M. K. & Martin, G. R. *Proc. Natl Acad. Sci. USA* **85**, 8998–9002 (1988).
5. Knight, C. A., Cheng, C. & DeVries, A. L. *Biophys. J.* **59**, 409–418 (1991).
6. Arrondo, J. L. R., Young, N. M. & Mantsch, H. H. *Biochim. Biophys. Acta* **952**, 261–268 (1988).
7. Li, N., Kendrick, B. S., Manning, M. C., Carpenter, J. F. & Duman, J. G. *Arch. Biochem. Biophys.* **360**, 25–32 (1998).
8. Worrall, D. *et al. Science* **282**, 115–117 (1998).
9. Knight, C. A., DeVries, A. L. & Oolman, L. D. *Nature* **308**, 295–296 (1984).
10. Duman, J. G., Wu, D. W., Olsen, T. M., Urrutia, M. & Tursman, D. *Adv. Low-Temp. Biol.* **2**, 131–182 (1993).
11. Graham, L. A., Liou, Y.-C., Walker, V. K. & Davies, P. L. *Nature* **388**, 727–728 (1997).
12. Li, N., Andorfer, C. A. & Duman, J. G. *J. Exp. Biol.* **201**, 2243–2251 (1998).
13. Hall, D. G. & Lips, A. *Langmuir* **15**, 1905–1912 (1999).

Cell differentiation

Hepatocytes from non-hepatic adult stem cells

Stem cells are undifferentiated long-lived cells that are capable of many rounds of division. Here we show that adult human liver cells can be derived from stem cells originating in the bone marrow or circulating outside the liver, raising the possibility that blood-system stem cells could be used clinically to generate hepatocytes for replacing damaged tissue.

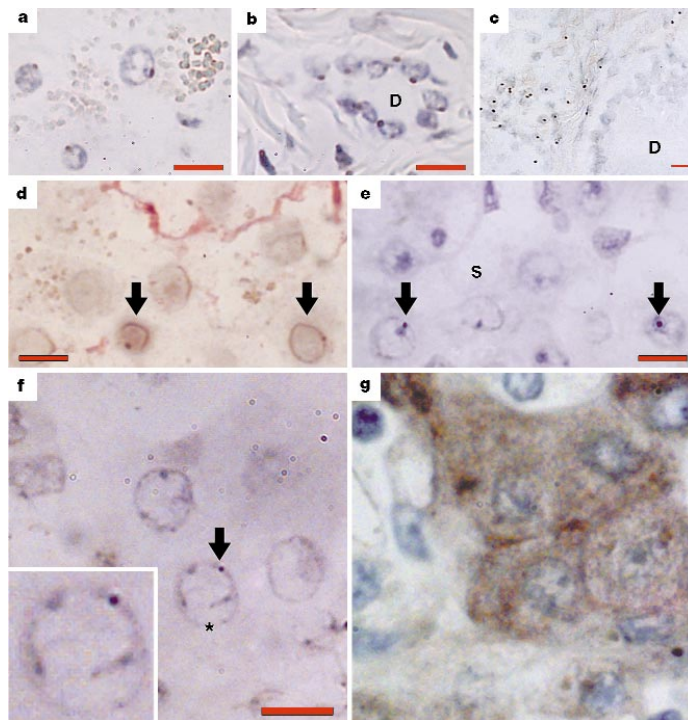
Serial transplantation in mice has indicated that some hepatocytes have the stem-cell-like property of massive division potential¹, suggesting that hepatocytes themselves are the principal progenitor cells for new hepatocytes. Also, when hepatocyte regeneration is compromised after injury, facultative stem cells in the bile ducts are activated, producing so-called oval cells that eventually differentiate into hepatocytes².

Cross-gender and whole-liver transplantation studies in rodents indicate that bone-marrow-derived or extrahepatic stem cells can differentiate into hepatocytes^{3,4}, so we investigated whether adult human haematopoietic stem cells could contribute to the regeneration of hepatocytes in damaged human liver tissue. First, we tested livers of female patients ($n=9$) who had received a bone-marrow transplant from a male donor for cells of donor origin by using a DNA probe specific for the centromere of the Y chromosome^{5,6}. Second, we looked for Y-chromosome-positive hepatocytes in female livers grafted into male patients ($n=11$) that were later removed because of recurrent disease. Y-chromosome-positive hepatocytes should indicate an extrahepatic origin for these cells in both cases.

Our probe detected the Y chromosome in most cells of all types in male control livers (Fig. 1a, b); as expected, inflammatory cells in the transplanted female liver in male recipients were Y-chromosome-positive and so served as a positive control (Fig. 1c). Protease digestion used as part of the Y-chromosome detection procedure caused some loss of morphology (Fig. 1d), but hepatocytes were still recognizable by their immunoexpression of cytokeratin 8 (ref. 7): many were Y-chromosome-positive, indicating that they originated from male-donor bone marrow.

It has been suggested that microchimaerism for Y-chromosome-positive cells occurs in the human female liver as a result of the transplacental passage of male fetal blood cells during pregnancy⁸. However, we did not detect any Y-chromosome-positive leukocytes or hepatocytes⁹ in liver biopsies taken from such women in a fluorescence *in situ* hybridization (FISH) assay. Although differentiation of fetal

Figure 1 Photomicrographs showing the presence of the Y chromosome, detected by immunolocalization of a fluorescein isothiocyanate (FITC)-labelled Y-chromosome probe using an anti-fluorescein antibody conjugated to horseradish peroxidase, and visualized using diaminobenzidine as a brown chromogen. The presence of the Y chromosome is indicated by a distinct brown dot, typically located at the nuclear periphery. **a**, Hepatocytes were identified by their large round nuclei and cytoplasmic granules and were Y-chromosome-positive in male controls. **b**, The Y chromosome was also detected in the bile ducts (D) in male controls. **c**, A Y-chromosome-negative bile duct surrounded by numerous Y-positive inflammatory cells in a female liver transplanted into a male. **d**, Immunodetection of two Y-chromosome-positive cells (arrows) located in a hepatocyte plate delineated by cytokeratin-8 immunostaining (red) from a female patient who had received male bone marrow. **e–g**, Y-chromosome detection in female livers transplanted into male recipients; **e**, two Y-chromosome-positive hepatocytes (arrows) in a hepatocyte plate bordering a sinusoid (S); note the brown dot, demonstrating that the Y chromosome is readily distinguishable from the blue nucleolus; **f, g**, consecutive sections of a group of four hepatocyte nuclei, showing **f**, one Y-chromosome-positive hepatocyte (arrow, and inset at 2-fold magnification), and **g**, their cytokeratin-8 immunoreactivity. Scale bars, 10 μ m.



bone-marrow stem cells into hepatocytes might occur after placental transfer, this is unlikely in the patient whose liver biopsy is shown in Fig. 1d because there was no history of male childbearing or abortion.

Chronic damage in the livers of our bone-marrow-transplant patients could promote the colonization and amplification of exogenous haematopoietic stem cells, providing a rationale for experimental models of hepatocyte transplantation in which blocking regeneration by indigenous cells or promoting their apoptosis encourages growth of transplanted cells in the damaged liver^{10,11}.

So do an individual's haematopoietic stem cells mobilize during liver failure to increase the regenerative capacity of their liver? We found examples of Y-chromosome-positive epithelial cells whose identity as hepatocytes was confirmed by their location and expression of cytokeratin 8 (Fig. 1e–g), indicating that circulating extrahepatic stem cells of endogenous, as well as exogenous, origin, can colonize the liver. Our liver samples were each from unique clinical cases, requiring individual protease digestion; the frequency of Y-chromosome-positive hepatocytes was relatively low in all of them (about 0.5–2%) and they often appeared as clusters (Fig. 1d, e), as though clonal growth had occurred after colonization.

Haematopoietic stem cells can be readily harvested from an individual, and we have shown that the human adult haematopoietic

stem-cell population is capable of being instructed by the appropriate environment to yield an epithelial lineage. Our results should contribute to the development of human tissue for use in a therapeutic context.

Malcolm R. Alison*, Richard Poulson†, Rosemary Jeffery†, Amar P. Dhillion‡, Alberto Quaglia‡, Joe Jacob*, Marco Novelli§, Grant Prentice||, Jill Williamson¶, Nicholas A. Wright*†

*Department of Histopathology, Imperial College School of Medicine, Hammersmith Campus, London W12 0NN, UK

e-mail: m.alison@ic.ac.uk

†Histopathology Unit, Imperial Cancer Research Fund, London WC2A 3PX, UK

‡Departments of Histopathology and

||Haematology, Royal Free Campus, Royal Free and University College School of Medicine, London NW3 2PF, UK

§Department of Histopathology, University College London, London WC1E 6JJ, UK

¶Human Cytogenetics Laboratory, Imperial Cancer Research Fund, London WC2A 3PX, UK

1. Overturf, K. *et al. Am. J. Pathol.* **151**, 1273–1280 (1997).

2. Alison, M. R. *Curr. Opin. Cell Biol.* **10**, 710–715 (1998).

3. Petersen, B. E. *et al. Science* **284**, 1168–1170 (1999).

4. Theise, N. D. *et al. Hepatology* **31**, 235–240 (2000).

5. Wolfe, J. *et al. J. Mol. Biol.* **182**, 477–485 (1985).

6. Novelli, M. R. *et al. Science* **272**, 1187–1190 (1996).

7. Van Eyken, P. & Desmet, V. J. *Liver* **13**, 113–122 (1993).

8. Tanaka, A. *et al. Hepatology* **30**, 833–838 (1999).

9. Rubbia-Brandt, L. *et al. Hepatology* **30**, 821–822 (1999).

10. Laconi, E. *et al. Am. J. Pathol.* **153**, 319–329 (1998).

11. Oren, R. *et al. Hepatology* **30**, 903–913 (1999).

Structure of the Fc fragment of human IgE bound to its high-affinity receptor FcεRIα

Scott C. Garman^{*†}, Beth A. Wurzburg^{*†}, Svetlana S. Tarchevskaya^{*}, Jean-Pierre Kinet[§] & Theodore S. Jardetzky^{*}

^{*} Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, Illinois 60208, USA

[§] Department of Pathology, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, Massachusetts 02215, USA

[†] These authors contributed equally to this work

The initiation of immunoglobulin-E (IgE)-mediated allergic responses requires the binding of IgE antibody to its high-affinity receptor, FcεRI. Crosslinking of FcεRI initiates an intracellular signal transduction cascade that triggers the release of mediators of the allergic response. The interaction of the crystallizable fragment (Fc) of IgE (IgE-Fc) with FcεRI is a key recognition event of this process and involves the extracellular domains of the FcεRI α-chain. To understand the structural basis for this interaction, we have solved the crystal structure of the human IgE-Fc–FcεRIα complex to 3.5-Å resolution. The crystal structure reveals that one receptor binds one dimeric IgE-Fc molecule asymmetrically through interactions at two sites, each involving one Cε3 domain of the IgE-Fc. The interaction of one receptor with the IgE-Fc blocks the binding of a second receptor, and features of this interaction are conserved in other members of the Fc receptor family. The structure suggests new approaches to inhibiting the binding of IgE to FcεRI for the treatment of allergy and asthma.

The high-affinity IgE receptor (FcεRI) is found on the surface of effector cells of the immune system that initiate cellular reactions associated with the allergic response, anaphylaxis and anti-parasitic immunity^{1,2}. The human receptor can form either a trimeric αγ₂ or tetrameric αβγ₂ structure on cell surfaces, with the extracellular domains of the α-chain conferring the ability to bind antibodies of the IgE class with high affinity ($K_d \approx 10^{-9}$ – 10^{-10} M). IgE antibodies are heterotetramers of two heavy and two light chains (H₂L₂), with the FcεRI-binding domains located in the carboxy-terminal Fc region of the heavy chain. The dimeric Fc fragment of the IgE antibody has full binding activity for FcεRI (refs 3–5). IgE antibodies bind to the receptor in the absence of antigen, and thus the receptor adopts the antigenic specificity of the prevalent IgE repertoire. Crosslinking of the receptor through antigen/antibody interactions leads to the initiation of a signal transduction cascade mediated by lyn and syk kinases, analogous to that induced by T-cell and B-cell receptors^{6–8}. In mast cells, receptor activation leads to rapid degranulation and the release of histamine, followed by the synthesis and release of prostaglandins, leukotrienes, cytokines and other mediators of the allergic response. Anti-parasitic responses can be triggered through a similar activation of eosinophils, leading to the release of granular proteins toxic to schistosomes and other parasites. FcεRI belongs to a family of antibody-binding receptors that also mediate interactions of soluble IgG antibodies with cells of the immune system^{6,8}.

Atopic diseases such as allergy, asthma and eczema comprise a wide spectrum of pathologies associated with the inappropriate activation of the immune system by environmental antigens⁹. Marked increases in atopic disease have been observed in the past century, particularly in developed countries. Genetic studies in both mice and humans have shown the association of the IgE network with allergic diseases, suggesting that polymorphisms of the FcεRI β-chain and CD14 are involved in atopic individuals⁹. The interaction of the IgE antibody with FcεRI is central to these immune reactions, and provides an attractive target for the inhibition of all IgE-mediated allergic disease. Clinical studies of allergic individuals using anti-IgE monoclonal antibody therapy has shown that this is an effective approach to disease treatment^{10,11}.

We have reported the crystal structure of the α-chain ectodomains of human FcεRI (ref. 12), which revealed a highly bent arrangement of two immunoglobulin domains. Four solvent-exposed tryptophans cluster at the top of the receptor, forming a large hydrophobic surface for potential interactions with the IgE-Fc. This tryptophan cluster borders the Fc-binding site that was previously mapped by mutagenesis studies, which implicated residues in the second domain of the receptor in IgE binding. The structural and functional data suggested that a large convex surface of the receptor could be involved in binding IgE, raising questions about the role of the tryptophans, the convex nature of the binding site, and the mechanisms underlying the stoichiometry and binding specificity for IgE. Given the dimeric structure of the IgE-Fc, which could potentially display two receptor binding sites, the structural basis for the observed stoichiometry of one IgE binding to one receptor has remained obscure.

To address these questions, we have solved the crystal structure of a complex of the human IgE-Fc bound to FcεRIα. The structure of the complex reveals two interaction sites for the IgE-Fc on the receptor surface and clarifies how a 1:1 complex between antibody and receptor is formed. The two Cε3 domains of IgE-Fc bind to distinct sites on the receptor: one is formed by the C–C' loop in the receptor D2 domain; the second involves the four solvent-exposed tryptophans. The structure of the complex accounts for previous mutagenesis and structural observations and shows that the Fc forms a complementary arch across the convex surface of the receptor. The IgE-Fc–FcεRIα complex provides a model for understanding the function of other antibody Fc receptors, and suggests new approaches to the inhibition of diseases mediated by IgE.

Structure determination of the complex

We expressed IgE-Fc and FcεRIα in insect cells using recombinant baculovirus technology. The expression of FcεRIα was carried out as described¹². The IgE heavy chain contains four constant domains (Cε1–Cε4), in contrast to the three found in IgG antibodies. The interaction of FcεRI with IgE has been mapped to the two C-terminal constant domains of the IgE-Fc (domains Cε3–Cε4)^{3,13–16}. Deglycosylated IgE-Fc retains comparable affinity for FcεRIα, but

aggregates into a less active form³. The treatment of cells that express IgE with oligosaccharide-processing inhibitors that limit glycosylation to a high-mannose form revealed no effects on FcεRI binding¹⁷.

The Cε3–Cε4 domains of human IgE–Fc isolated from insect cells are modified with high-mannose carbohydrate at the conserved glycosylation site at N394 (B.A.W. *et al.*, manuscript in preparation), and form stable complexes with FcεRIα. Using wild-type FcεRIα protein, crystals of the complex grow in space group *P*₄₁₂₁₂; however, diffraction data are limited to a resolution of about 4.5 Å (Table 1, crystal form I). To improve the crystals, we expressed a triple carbohydrate mutant of FcεRIα (FcεRIαΔ4–6) in insect cells. The FcεRIαΔ4–6 mutant lacks carbohydrate attachment sites at three of the seven wild-type locations (residues 74, 135 and 140)¹⁸. Complexes formed with FcεRIαΔ4–6 expressed in baculovirus grow crystals in space group *R*32 and diffract X-rays to a resolution of 3.5 Å (Table 1, crystal form II). The structure was determined by molecular replacement techniques (see Methods). Current refinement statistics for the complex are shown in Table 1, with an overall *R*_{free} of 27.3% and *R*_{work} of 25.4% to 3.5 Å. Figure 1a illustrates electron density from a σ_a -weighted $2F_o - F_c$ simulated annealing omit map calculated with the current model phases. Both crystal forms of the IgE–Fc–FcεRIα complex show a 1:1 complex with similar overall features. Interpretation of the structure is limited to crystal form II because of the low resolution of crystal form I.

Overview of the complex

The receptor contains two distinct binding sites for the antibody (Fig. 1b–d). Binding interactions are formed exclusively between the Cε3 domains of the IgE–Fc and FcεRIα. The Cε3–Cε4 domain interfaces and the Cε4 domains do not make receptor contacts. The two Cε3 domains are related by a nearly perfect dyad axis (180.7° rotation), except for residues in the Cε2–Cε3 linker region (residues 328–336). The Cε4 domains are also related by a nearly perfect dyad axis (179.6° rotation), which is offset 3° from the Cε3 pseudo-dyad (see below). Electron density for the terminal residues of each chain is weak, including the amino-terminal interchain disulphide of the IgE–Fc.

There are *N*-linked oligosaccharides in both the IgE–Fc and FcεRIα (Fig. 1b–d), but they do not contribute significantly to interactions between the two molecules. Carbohydrate attached to the conserved site (N394 in IgE), lies along the inner face of the Cε3

domains (strands A, B and E) similar to that observed in structures of IgG–Fc. Carbohydrate attached to residue N42 in FcεRIα extends upwards to the top of the D1–D2 interface approaching one of the IgE–Fc chains (Fig. 1b–d). Carbohydrate residues are attached to N21 and N166 of FcεRIα, projecting from the receptor surface away from the IgE-binding interface. We did not model carbohydrate for N50 of FcεRIα, nor N371 or N383 of IgE.

The Cε3 domains of the heavy chain dimer of IgE bind along one side of the D2 domain and at the top of the D1–D2 interface of FcεRIα, defining two distinct interaction sites. Site 1 refers to the interaction of one Cε3 domain with the C–C' region of the receptor D2 domain, centred around receptor residue Y131. Site 2 refers to the interaction of the second Cε3 domain with the D1–D2 interface (Fig. 1b, c), and involves the cluster of four surface-exposed tryptophans (W87, W110, W113 and W156). The two chains of the IgE–Fc bind the receptor using surface loops in Cε3 (Fig. 2a), including the immunoglobulin-fold of BC loop (362–365), DE loop (393–396), FG loop (424–427), and the Cε2–Cε3 linker region (332–336). Interactions with FcεRIα cause the Cε2–Cε3 linker regions to point up and away from the complex interface. The Cε4 domains of the IgE–Fc provide a structural dimerization scaffold that allows two Cε3 domains to form the bivalent interaction with FcεRIα.

Structural basis for the formation of a 1:1 complex

Biophysical studies of the IgE–Fc–FcεRIα complex indicate that a 1:1 complex is formed between the antibody and receptor^{4,5,19,20}, consistent with the crystallographic structure. This contrasts with models proposing a 1:2 stoichiometry for the homologous IgG–Fc receptors FcγRIIa²¹ and FcγRIIb^{22,23}. Fc receptors belonging to other structural classes can form complexes with a 1:2 stoichiometry, such as the major histocompatibility complex (MHC)-like neonatal Fc receptor²⁴, or use two identical domains to bind one antibody, such as the lectin-like low-affinity IgE receptor FcεRII (ref. 25). The 1:1 stoichiometry of the IgE–Fc–FcεRIα complex ensures that receptor crosslinking and cellular activation depend upon multivalent antigen binding.

Figures 1b–d and 2 show surface and ribbon representations of the IgE–Fc–FcεRIα complex. The two Cε3 domains of IgE interact with FcεRIα using surface loops that are structurally analogous to the complementarity determining region loops of antibody variable

Table 1 Data collection and refinement statistics

Data	Form I	Form II, low resolution	Form II, high resolution
Data Set			
Resolution (Å) *	30–4.50 (4.66–4.50)	30–4.00 (4.14–4.00)	40–3.5 (3.63–3.5)
Source	APS DND 5ID	ALS 5.0.2	APS DND 5ID
Wavelength (Å)	1.000	1.200	1.034
Completeness*	99.5 (97.7)	99.7 (98.3)	99.8 (99.7)
Average redundancy*	7.0 (5.8)	5.0 (3.9)	3.1 (3.1)
<i>R</i> _{sym} (%)†	17.8 (57.5)	15.2 (75.0)	9.1 (46.1)
$\langle I/\sigma \rangle$ *	5.9 (2.0)	4.4 (2.0)	16.5 (2.7)
No. of observed / no. of unique reflections	39,925 / 5,703	91,617 / 18,459	84,873 / 27,484
No. of working / no. of free *		18,455 / 945	25,993 / 1,418 (2,543 / 159)
<i>R</i> _{work} / <i>R</i> _{free} (%)‡			25.4 / 27.3 (31.6 / 37.2)
Refinement (form II, 3.5 Å)			
No. of atoms	Total	Protein	Carbohydrate
	5,251	4,821	259
		Detergent	Sulphate
		146	25
		Water	0
R.m.s. Deviations	Bonds	Angles	Dihedrals
	0.0100 Å	1.55°	25.4°
			Improper
			1.23°
Average <i>B</i> (Å ²)	Overall	Receptor	Fc chain 1
	88.0	68.1	97.8
			Fc chain 2
			99.2
Ramachandran	Favoured	Allowed	Generous
	77.0%	21.5%	1.5%
			Disallowed
			0.0%

* Last shell is shown in parentheses.

† $R_{\text{sym}} = \sum_i \sum_j |I_{ij} - \langle I_i \rangle| / \sum_i \sum_j I_{ij}$, where I_{ij} is the i^{th} intensity measurement of reflection h and $\langle I_i \rangle$ is the average intensity of that reflection.

‡ $R_{\text{work}}/R_{\text{free}} = \sum_h |F_o - F_c| / \sum_h |F_o|$, where F_c is the calculated and F_o is the observed structure factor amplitude of reflection h for the working/free set, respectively.

domains (Fig. 2a). The two distinct binding surfaces of the receptor (Fig. 2b) establish a convex and asymmetric interaction with the two $\text{Ce}3$ domains, placing the receptor near the $\text{Ce}3$ pseudo-dyad axis. There are two structural keys that dictate the formation of complexes with 1:1 stoichiometry: first, induction of structural asymmetry in the $\text{Ce}2$ – $\text{Ce}3$ linker of IgE-Fc; and second, steric hindrance that blocks the binding of a second receptor.

Superimposing the two $\text{Ce}3$ domains in the IgE-Fc structure (Fig. 2c) shows that the $\text{Ce}2$ – $\text{Ce}3$ linker regions (residues 328–336) are constrained to an asymmetric arrangement by interactions with $\text{Fc}\epsilon\text{RI}\alpha$. The $\text{Ce}4$ domains are not optimally aligned by the superposition of the $\text{Ce}3$ domains, owing to the 3° difference in the $\text{Ce}3$ and $\text{Ce}4$ pseudo-dyad axes. Independent movement of $\text{Ce}3$ and $\text{Ce}4$ domains within one heavy chain might allow optimal binding interactions to form in the complex. The BC and FG loops also adopt slightly different conformations in the two $\text{Ce}3$ domains because of their interactions with the two binding sites of $\text{Fc}\epsilon\text{RI}\alpha$.

Binding of one receptor to the IgE-Fc creates a steric block to binding a second receptor. Figure 2d–f shows representations of both the IgE-Fc and $\text{Fc}\epsilon\text{RI}\alpha$ in which the complex has been separated to expose the buried interaction surfaces. The amino acids of the $\text{Ce}2$ – $\text{Ce}3$ linker form the top of an arch that conforms to the convex surface of $\text{Fc}\epsilon\text{RI}\alpha$, creating an asymmetric binding site for a single receptor. In addition, four IgE residues (R334, G335, V336 and H424) are common to both site 1 and site 2, preventing the simultaneous binding of two receptors to one IgE-Fc. Superposition of a second receptor onto the 1:1 complex shows significant steric overlap between receptors and the amino acids of the $\text{Ce}2$ – $\text{Ce}3$ linker. Thus, the binding of one receptor effectively prevents the binding of a second, owing to both the induced asymmetry of the $\text{Ce}2$ – $\text{Ce}3$ linker of the IgE-Fc and the steric inhibition across the Fc dyad axis.

Structural changes in the receptor and IgE on binding

The receptor shows little change in conformation on complex formation with the Fc. The overall r.m.s. difference in 158 $\text{C}\alpha$ positions is 1.11 Å compared with the unbound receptor, with little change in the angle between D1 and D2 (ref. 12). Two loops on the receptor adopt conformations different from those in the uncomplexed $\text{Fc}\epsilon\text{RI}\alpha$ structure, the C' strand in D2 (residues 127–133) and the BC loop in D1 (residues 30–35). The C' strand of D2 is longer in the complex; residues L127–Y131 form additional hydrogen bonds to the C strand. Analysis of the $\text{Fc}\epsilon\text{RI}\alpha$ structure in several crystal forms (S. C. G. *et al.*, manuscript in preparation), however, shows that the C' strand can adopt this conformation in the absence of the IgE-Fc. The BC loop in D1 also adopts different conformations in different crystal forms, but this region is distal to the IgE-Fc binding site.

In the complex, the IgE-Fc is in a conformation that is similar to the Fc domains of IgG antibodies^{26,27}, but the structure of the uncomplexed IgE-Fc reveals structural differences that could influence receptor binding (B. A. W. *et al.*, manuscript in preparation).

Binding surfaces of the IgE-Fc– $\text{Fc}\epsilon\text{RI}\alpha$ interaction

Formation of the IgE-Fc– $\text{Fc}\epsilon\text{RI}\alpha$ complex buries a binding surface totalling about 1,850 Å² (Fig. 2d–f). Site 1 and site 2 use overlapping but non-identical sets of IgE residues. Of the fifteen $\text{Fc}\epsilon\text{RI}\alpha$ residues within 4 Å of the IgE-Fc, seven are aromatic—of which five are surface-exposed tryptophans. In contrast, of the eighteen IgE-Fc residues within 4 Å of $\text{Fc}\epsilon\text{RI}\alpha$, none are aromatic. The large number of aromatic receptor residues and the large buried surface area may contribute to the stability of the complex ($K_d \approx 10^{-9}$ – 10^{-10} M).

Figure 3a shows a plot of the IgE-Fc residues that are buried by interaction with the receptor. Residues of the $\text{Ce}3$ domain located in site 1 (top half of Fig. 3a) form specific interactions with the residues

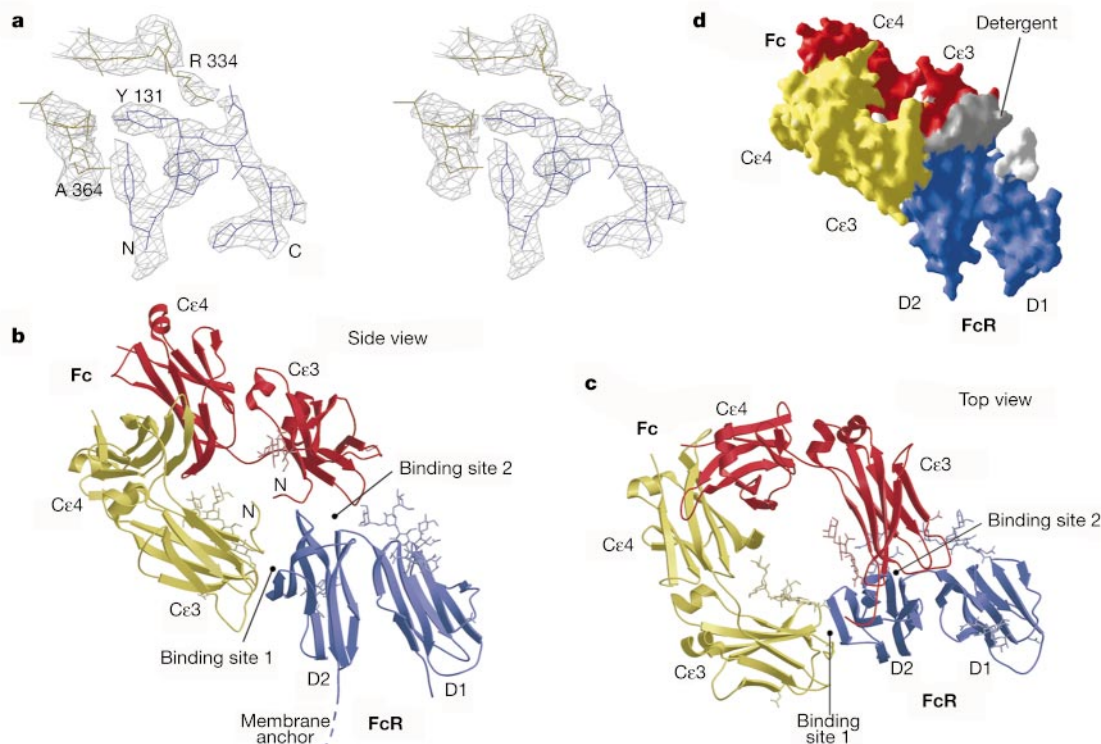


Figure 1 Overview of the crystal structure of the IgE-Fc– $\text{Fc}\epsilon\text{RI}\alpha$ complex. **a**, Example electron density for binding site 1. Shown is a stereo diagram from a σ_A -weighted $2F_o - F_c$ simulated annealing omit electron-density map at 3.5 Å calculated using the final refined structure of the complex contoured at 1.25 σ . Residues 129–136 of $\text{Fc}\epsilon\text{RI}\alpha$, and loop residues 334–336 and 362–364 of IgE-Fc are shown. **b**, Side view of the IgE-Fc– $\text{Fc}\epsilon\text{RI}\alpha$ complex. The two Fc chains are in yellow and red, the $\text{Fc}\epsilon\text{RI}\alpha$ chain is in blue.

Carbohydrate residues are shown as sticks in the same colour as the protein chain to which they are attached. The colour scheme is the same in all figures. Binding sites 1 and 2 are indicated. The cell membrane would lie below the receptor. **c**, Top view of the IgE-Fc– $\text{Fc}\epsilon\text{RI}\alpha$ complex. **d**, Side view of a surface representation of the IgE-Fc– $\text{Fc}\epsilon\text{RI}\alpha$ complex highlighting the asymmetric interactions on the receptor surface. Carbohydrate surfaces are white, detergent surface is grey.

of FcεRIα shown in Fig. 3b. Twelve amino acids from the IgE and eight amino acids from the receptor form site 1, burying a total of about 860 Å² of surface area. The IgE residues are from four distinct regions of the IgE-Fc sequence, including the Cε2–Cε3 linker (residues 334–336), the BC loop (362–365), the DE loop (393–396) and the FG loop (residue 424). The first *N*-acetylglucosamine attached at the conserved N394 site buries 50 Å² in site 1. N394 and the attached carbohydrate are not involved in site 2. The limited involvement of the carbohydrate is consistent with biochemical studies showing that it is not required for receptor binding³. The receptor residues derive from two regions of the D2 domain, involving the C strand (residues 117, 119 and 121) and the C'–E region (residues 126 and 129–132). Two potential salt bridges (αK117–Cε3D362 and αE132–Cε3R334) and four potential hydrogen bonds (αK117–Cε3G335, αY129–Cε3D362, αY131–Cε3D364 and αY131–Cε3H424) are formed across the site 1 interface (Fig. 3b). Y131 from the receptor projects into a pocket

on the IgE-Fc formed by the BC and FG loops of Cε3 and the Cε2–Cε3 linker (Figs 2d–f and 3b).

The residues the Cε3 domain that are buried in the formation of site 2 are shown in Fig. 3a and c. Interactions at site 2 bury about 970 Å², including ten amino acids from IgE and eight from FcεRIα. The IgE residues are localized to two distinct segments of Cε3: the Cε2–Cε3 linker region (residues 332–337); and the FG loop (424–427). The FcεRIα contributes residues from three regions: the D1–D2 linker region (85–87); the D2 BC loop (residues 110 and 113); and the D2 FG loop (156–158). Residues from the D1 domain of the receptor do not form direct interactions with the IgE-Fc. Site 2 primarily contains hydrophobic amino acids with three potential hydrogen bonds across the interface (αW156–Cε3G335, αQ157–Cε3N332 and αQ157–Cε3R334). P426 from the IgE-Fc intercalates between receptor residues W87 and W110, forming a hydrophobic proline sandwich (Figs 2d, e and 3c). These three residues are conserved in IgG antibodies and the IgG receptor family.

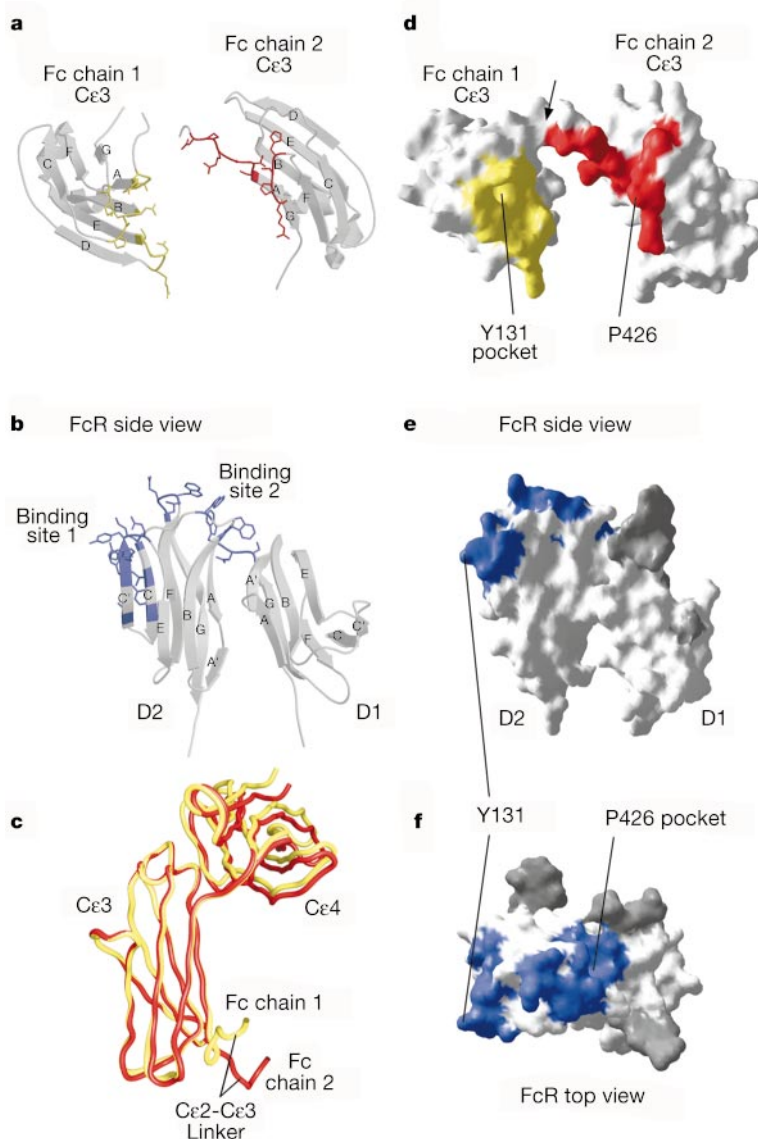


Figure 2 Structural basis for the 1:1 binding of IgE to FcεRIα. The IgE-Fc–FcεRIα complex has been taken apart to display the interaction sites. **a**, End-on view of the Cε3 domains of IgE. Immunoglobulin domain strands are labelled, and binding residues that bind FcεRIα are shown as sticks in yellow (site 1) and red (site 2). **b**, Side view of FcεRIα. Strands in D1 and D2 are labelled, and residues that bind IgE are shown as sticks on the ribbon surface. **c**, Superposition of the two Cε3 domains of IgE-Fc. The twofold symmetry of the IgE-Fc domains is broken in the Cε2–Cε3 linker region (328–336) by interactions

with the receptor. Superposition of the Cε3 domains leads to a small displacement in the Cε4 domains, because of a 3° difference in Cε3 and Cε4 pseudo-dyad axes. **d**, Surface view of the Cε3 domains of IgE in the same orientation as in **a**. The surface buried by FcεRIα is coloured yellow (site 1) or red (site 2). The binding pocket for receptor residue Y131 is shown, as is the conserved P426 on the Fc. An arrow indicates the Cε2–Cε3 linker. **e, f**, FcεRIα residues buried upon formation of the complex. Residue Y131 of site 1 and the binding pocket for P426 of the IgE-Fc are labelled. Carbohydrate is shown in grey.

Electron density is observed for CHAPS detergent in the form II crystals, with one molecule located above W156 of FcεRIα and below the FG loop of Ce3 near H424 in site 2 (Fig. 3d). Although the CHAPS interaction may be weak, this structure provides a foundation for using combinatorial synthetic chemistry methods to improve these initial binding interactions²⁸. The position of the CHAPS steroid core is analogous to the position of Y131 of FcεRIα in site 1, and adding atoms to the CHAPS scaffold could disrupt the interactions in site 2. H424 of the IgE-Fc, which is located next to the CHAPS binding site, makes contacts with the receptor in both site 1 and site 2. A small molecule inhibitor that could interact with both the Y131 pocket of Ce3 (site 1) and with H424 might effectively disrupt both site 1 and site 2 interactions with the receptor.

Locations of IgE and FcεRIα mutations

Mutagenesis studies of both the IgE-Fc and FcεRIα have been carried out in efforts to define the residues contributing to the stability of the complex. For FcεRIα, these studies have implicated residues located in the D2 domain, including residues W87, W113, V115, K117, V118, Y120, Y121, K122, D123, K128, Y129, W130, Y131, E132, Y149, G153, V155, W156, D159, Y160 and E161 (refs 12, 29–33). Not all of these make direct contacts (within 4 Å) with the IgE-Fc. Of the residues identified by mutagenesis, eight interact directly with the IgE (W87, W113, K117, Y129, W130, Y131, E132 and W156), twelve are within three residues (V115, V118, Y120, Y121, K122, D123, K128, G153, V155, D159, Y160 and E161), and the remaining one (Y149) is buried in the hydrophobic core of D2.

The identification of the FcεRIα-binding site in IgE-Fc has implicated many regions in the Ce3 domain, including the Ce2–Ce3 linker, AB helix, CD loop, DE loop, EF helix, FG loop and G strand^{3,13,15,16,34}. Mutagenesis studies have identified residues P333, R334, F349, R376, S378, K380, R393, D409, E414, R427 and M430

as possible contact residues in the IgE-Fc. Of these residues, three are observed as contact residues (P333, R334 and R427) and one (M430) is within three residues of a contact. Six are distant from the IgE-Fc–FcεRIα interface (F349, R376, S378, K380, D409 and E414). Some mutations of these distant residues have little effect on FcεRIα binding (such as R376A, R376K, D409A, D409E and D409N), whereas others reduce binding to the receptor (such as R376E and D409R). Mutations distant from the FcεRIα binding site (in the AB helix, CD loop and EF helix) may indirectly affect receptor binding by altering the Ce3 conformation or the interdomain angle between Ce3 and Ce4.

The basis for IgE specificity and its implications

Diagrams of the amino-acid residues that lie within 4 Å of each other in site 1 and in site 2 are shown in Fig. 4a and b. Connecting lines highlight residues forming atomic contacts across the respective interfaces. Also shown are the residues that are found in the related human IgG receptors and in four subtypes of human IgG antibodies.

In site 1 there is little conservation of the residues that form the IgE-Fc–FcεRIα interface. Three residues are completely conserved (G335, D362 and N394 of IgE) in the Fc sequences, but there is poor conservation in the receptor sequences, except for the partial conservation of K117 and the relatively conserved Y129 (either tyrosine or phenylalanine). The conservation of K117 in three of the four receptors matches the complete conservation of D362 and G335 in the Fc sequences, potentially preserving one of the two salt bridges and one of the hydrogen bonds of site 1. For FcεRIα Y129, only tyrosine or phenylalanine is found, suggesting that this interaction may be conserved in IgG-Fc complexes with the FcγRs. However, Y131 of FcεRIα, which forms a large number of atomic contacts across the interface and is buried in a shallow surface

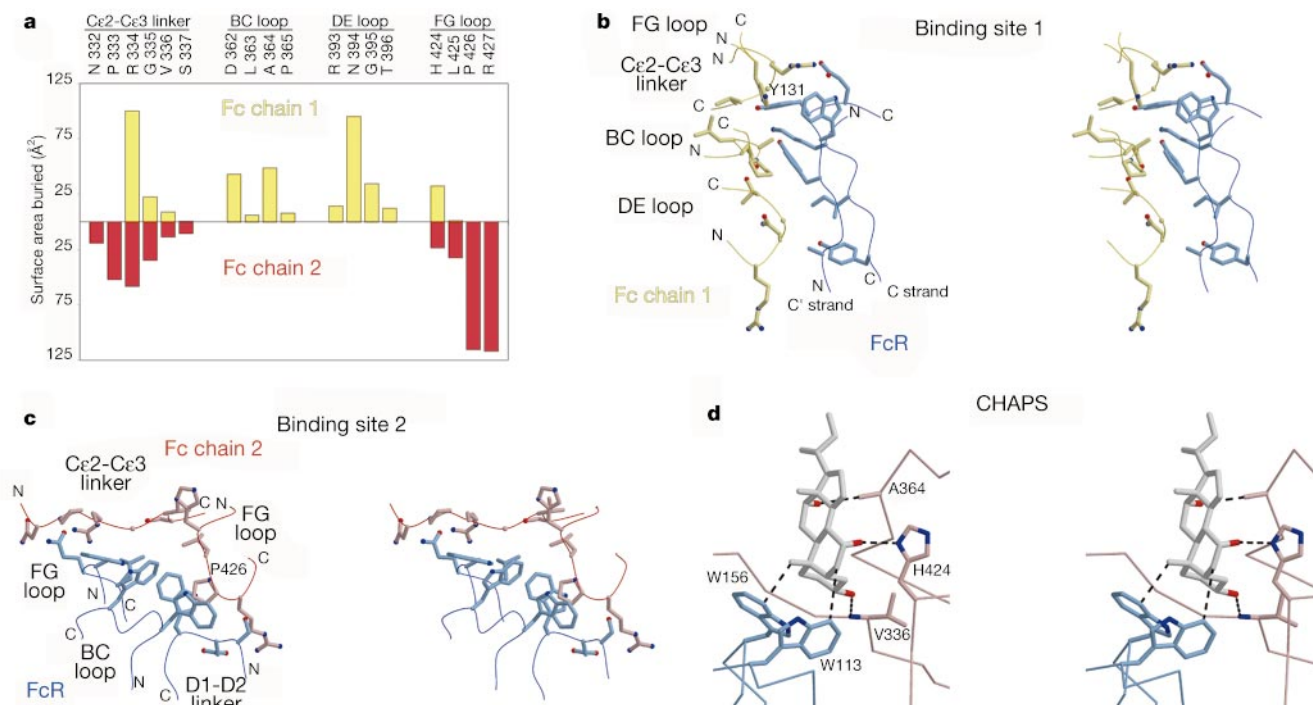


Figure 3 Details of the interactions in the IgE-Fc–FcεRIα complex at site 1 and site 2. **a**, Plot of the buried surface area of residues in the Ce3 domains of IgE-Fc. Residues buried in the site 1 interaction (top half; yellow bars) and residues buried in the site 2 interaction (bottom half; red bars) are shown. The IgE loops are identified above. Attached carbohydrate contributes 50 Å² of the buried surface area of N394. **b**, Stereo view of residue interactions at site 1. Binding loops are labelled at their termini, side chains of residues buried in the complex are shown, and Y131 is labelled. IgE-Fc and FcεRIα chains are yellow and blue, respectively. **c**, Stereo view of residue interactions at site 2.

Side chains of residues buried in the complex are shown. IgE-Fc and FcεRIα chains are red and blue, respectively. P426 (labelled) lies between the side chains of W87 and W110. **d**, Stereo view showing the binding of a CHAPS detergent molecule in the complex. Nitrogens are blue, oxygens are red, Fc carbons are pink, FcεRIα carbons are cyan and detergent carbons are white. Atoms less than 4 Å apart have dotted lines between them, and the residues are labelled. No density appears for the flexible end of the detergent, and those atoms are not modelled.

pocket on the IgE-Fc, is not conserved in the FcγRs (changing to either histidine or arginine). Given the central location of Y131 in the site 1 interface (Fig. 2d–f), this residue may generally be important in immunoglobulin specificity. Four of the five residues that contact Y131 in IgE change in the IgG-Fc sequences. Two allelic forms of FcγRIIa differ in only two residues (27 and 131) and in their binding to IgG2 (ref. 35). Residue 27 is not involved in the IgE-Fc–FcεRIα interface, but the change from R131 to H131 in site 1 may explain the loss of IgG2 binding for the R131 allotype. Residues in the four IgG subtypes are generally highly conserved in the site 1 interface (7 out of 9 are identical) compared with the significant variation in the FcγR residues.

Figure 4b shows the conservation of interactions in the site 2 interface. In the antibody sequences, P426 and L425 are absolutely conserved and P426 interacts with two absolutely conserved tryptophans in FcεRIα (W87 and W110), which form a hydrophobic pocket (Figs 2d, e and 3c). Site 2 also includes three residues (residues 332–334 of IgE) that in IgG affect binding to FcγRI. IgG1 binds with high affinity to FcγRI, whereas IgG2 does not. The binding affinity of IgG2 can be introduced into IgG1 by the substitution of residues LLG to PVA (in black in Fig. 4b), and IgG1 binding affinity is introduced into IgG2 by the reciprocal substitution^{36,37}. These residues in the IgE-Fc (residues 332–334) interact with the FG loop of FcεRIα (Figs 3c and 4b). Mutagenesis experiments have also shown that the transfer of the FG loop of FcεRIα to FcγRII confers detectable IgE binding³⁸. For site 2, the conservation of five residues and the additional functional data suggest a conserved binding mode across these members of the FcR

family. Variation in the FcγR FG loop sequences that contact the N-terminal linker region of the Fc fragment may provide key interactions that determine the specificity of FcγR–IgG pairs.

Discussion

The identification of two distinct binding sites in the IgE-Fc–FcεRIα complex and observed flexibility in Fc structures²⁷ (B. A. W. *et al.*, manuscript in preparation) suggest a kinetic scheme for IgE binding (Fig. 5a). Interdomain flexibility in the Cε3–Cε4 elbow could allow each Cε3 domain to transiently dissociate in the complex, leading to two pathways for the full dissociation of the IgE. Surface plasmon resonance studies of IgE-Fc–FcεRIα complexes show biphasic dissociation rates consistent with two distinct binding interactions^{16,34}. In this kinetic scheme, inhibitors specific for either site 1 or site 2 might accelerate the dissociation of receptor-bound IgE by capitalizing on the transient exposure of each site in the complex. Such inhibitors might prove useful in the treatment of acute allergic reactions in which dissociation of IgE from mast cells would be beneficial.

IgG-Fc and Fcγ receptors are structurally similar to the IgE-Fc and FcεRIα in the complex, and potentially could form analogous interactions. Superposition of a Cγ2 domain of IgG on the homologous Cε3 domain of IgE involved in binding site 2 places the second Cγ2 domain of IgG in close proximity to binding site 1 (Fig. 5b). Members of the Fc-receptor family have structural similarity, as shown by the superposition of the low-affinity IgG receptor FcγRIIb²² onto FcεRIα (Fig. 5b). Binding site 2 involves hydrophobic residues conserved in IgG and IgE antibodies and receptors, whereas binding site 1 and the receptor FG loop generate specificity through amino-acid variability. Compared with IgE, the binding of IgG to its receptors is more sensitive to the presence of N-linked carbohydrate in the Fc³⁹.

A model for an intact IgE antibody binding to FcεRIα is shown in Fig. 5c. In IgE, the Fc contains an additional pair of domains (Cε2) that separate the receptor-binding and Fab regions of the antibody. Calorimetric and binding studies show that the presence or absence

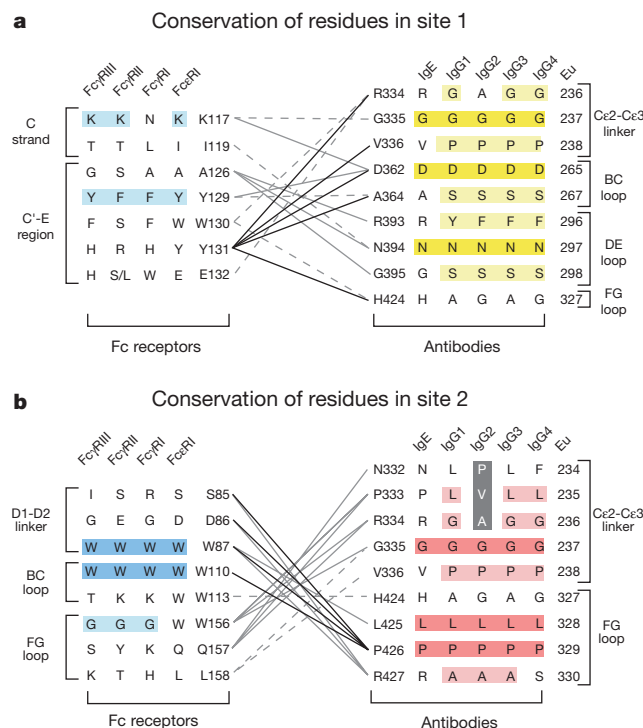


Figure 4 Conservation of amino-acid residues and contacts at the IgE-Fc–FcεRIα interfaces in IgG receptors and antibodies. Contacting residues are defined as interatomic distances less than 4 Å. **a**, Residues that interact in site 1, and their conservation in related human receptors and antibodies. Absolutely conserved residues are highlighted in bold, and partially conserved residues are lightly highlighted (yellow for IgE, blue for FcεRIα). Dark lines are drawn for residues making the largest number of contacts across the interface, lighter lines for intermediate numbers of contacts, and dashed lines for the fewest contacts. Numbering for the Eu protein (human myeloma IgG1) sequence is shown to the right. **b**, Residues that interact in site 2, and their conservation in related human Fc receptors and antibodies. Receptor residues are highlighted in blue, antibody residues in red. Three residues in IgG2 (P, V, A) that affect binding to FcγRI are boxed in black.

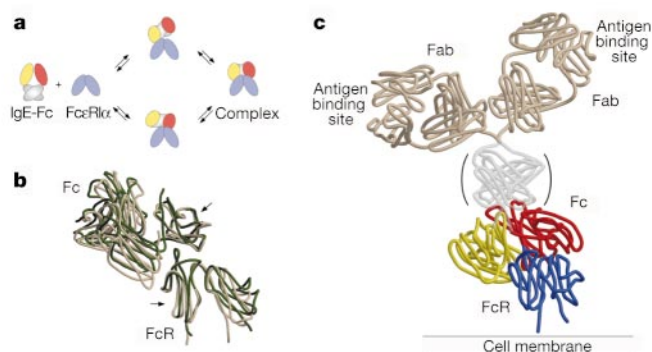


Figure 5 Implications of the complex. **a**, Kinetic scheme for the binding of IgE to its receptor. The interaction of each Cε3 domain with distinct surfaces of the FcεRIα structure suggests a kinetic scheme in which transient release of one of the Cε3 domains may occur within the complex. This could lead to two distinct pathways for the association and dissociation of the complex, consistent with the experimental observation of two distinct kinetic off-rates. Transient opening of the complex may allow inhibitors to enhance the dissociation of receptor-bound IgE by preventing the rebinding of an exposed Cε3 domain within the complex. **b**, Superposition of the Fc portion of an intact IgG antibody (1IGY)²⁷ and IgG Fc receptor FcγRIIb²² onto the IgE-Fc–FcεRIα complex. The IgE complex is shown in beige, the IgG homologues in green. Domains used to superimpose the structures are indicated with arrows. Only a minor adjustment of the other IgG domains is required to fit the IgE complex. **c**, A hypothetical model for an intact IgE-Fc–receptor complex. The IgE-Fc–FcεRIα complex is coloured as in Fig. 1, and antibody Fab regions are beige. The Cε2 domains of IgE are represented as paired immunoglobulin domains (transparent) inserted into the hinge regions between Fab and Fc, and shown in parentheses because no structural information is available for these domains.

of the Ce2 domains has no effect on the thermodynamic parameters associated with FcεRIα binding⁴⁰. In the absence of receptor binding, the Ce2 domains are probably tethered by flexible polypeptide segments, with no influence on the Ce3–Ce4 domain structure of IgE–Fc or on direct interactions with FcεRIα. In this model, Ce2 domains extend away from the receptor surface in the same direction as the Ce2–Ce3 linkers in the crystal structure, acting as an additional spacer between the Fabs and the Fc–FcR interface. Receptor binding at the Ce2–Ce3 linker promotes an overall bend to the IgE, consistent with previous fluorescence energy transfer studies⁴¹. Antibody bending would place the Fab regions away from the membrane (Fig. 5c), in position to bind multivalent antigens.

In contrast to receptors that dimerize and signal after binding a well defined ligand, Fc receptors act as adaptors between highly variable antigen-binding sites and the intracellular signal transduction machinery. Flexible linkers between the antibody Fab and Fc arms seem necessary to accommodate the variable structures of Fab–antigen complexes and the need to crosslink two or more Fc receptors to initiate signalling^{1,2,7}. Thus, FcR signal-inducing ligands (antigens) can promote co-localization of receptors but cannot form specific contacts that dimerize FcRs, as observed for hormone or cytokine receptor complexes. It remains unclear whether co-localization leads to a specific signalling structure or simply membrane relocation of the complex². Given the heavily glycosylated state of the FcεRIα chain, it is unlikely that the D1–D2 domains could form protein–protein contacts to initiate signalling. Contacts between neighbouring antibodies, α-chain transmembrane segments, or FcεRI β- or γ-subunits, however, might mediate the assembly of a distinct signalling complex after antigen-induced multimerization.

The crystal structure of the IgE–Fc–FcεRIα complex clarifies the atomic interactions that regulate the specificity and stoichiometry of protein–protein interactions underlying allergic reactions. The observation of bound CHAPS molecules provides a lead for the structure-based development of novel inhibitors for the treatment of allergy, anaphylaxis and asthma. Insights into antibody binding to Fc receptors may also provide new approaches to treating autoimmune disorders and to the improvement of cancer therapies based on antibodies. □

Methods

Crystallization of the human IgE–Fc–FcεRIα complex

We expressed the Ce3–Ce4 domains of human IgE–Fc (B. A. W. *et al.* manuscript in preparation) and FcεRIα¹² in insect cells. The purified IgE–Fc protein includes four non-wild-type residues generated by the construct (ADPC), followed by wild-type amino acids 330–547 (ref. 42). The cysteine (construct residue 329) is displaced by one amino acid from the wild-type IgE–Fc C328, and forms an interchain disulphide bond in about 95% of the isolated IgE–Fc. Complexes of the Ce3–Ce4 domains of IgE and wild-type FcεRIα produced poorly diffracting crystals (Table 1, crystal form I). As the receptor is heavily glycosylated (~33% carbohydrate by weight), and the carbohydrate sites are dispersed on the receptor surface, we selectively removed a subset of these attachment sites to improve the protein crystallization. We expressed a carbohydrate mutant of the receptor lacking three of the seven wild-type carbohydrate sites (74, 135 and 140)¹⁸ in insect cells. The receptor triple mutant FcεRIαΔ4–6 was purified by affinity chromatography¹².

We incubated purified IgE–Fc and FcεRIαΔ4–6 or FcεRIα to form a complex, and purified it by gel-filtration chromatography using a Superdex 75 column (Pharmacia). The complex was concentrated to 10 mg ml⁻¹, and hanging drop crystallizations carried out at 23 °C. Form I crystals of the wild-type IgE–Fc–wild-type FcεRIα complex grew from 1.4 to 1.6 M ammonium sulphate and 100 mM Tris pH 8.5, over a period of 8–12 months. Form II crystals grew from purified wild-type IgE–Fc–FcεRIαΔ4–6 complex using 100 mM Tris, pH 8.5, 1.4–1.6 M ammonium sulphate and 8 mM 3-[(3-cholamidopropyl)dimethylammonio]-1-propane-sulphonate (CHAPS) in several days. Neither crystal form grew longer than about 110 μm. Crystals were transferred to buffer (form I, 2.1–2.7 M ammonium sulphate and 100 mM Tris pH 8.5; or form II, 1.6–2.0 M ammonium sulphate, 100 mM Tris pH 8.5 and 0.8 mM CHAPS) and frozen in the same buffer supplemented with 15% glycerol. Data sets were collected at ALS 5.0.2 beamline and the APS DND-CAT 5-ID-B beamline at ~160 °C using an ADSC Quantum 4 detector or a MarCCD detector. Significant crystal decay was observed during data collection. Images were processed using the HKL suite of programs⁴³, and intensities were adjusted using TRUNCATE⁴⁴. Form I crystals belong to space group P4₁2₁2 with $a = b = 126 \text{ Å}$, $c = 129 \text{ Å}$, and form II crystals belong to the space group R32 with $a = b = 192.8 \text{ Å}$ and $c = 302.4 \text{ Å}$

(hexagonal setting). Both crystal forms contain one IgE–Fc–FcεRIα complex in the asymmetric unit.

Crystal structure determination and refinement

Using the AMoRe⁴⁴ and EPMR⁴⁵ programs, we carried out molecular replacement on the form II low-resolution data with model coordinates from the 2.4 Å structure of the receptor¹². Normalized structure factors were critical to the success of the search in AMoRe. Both AMoRe and EPMR produced crystallographically equivalent locations for the receptor. 2F_o–F_c electron density maps with phases from the receptor revealed density corresponding to the two Ce3 domains of the Fc. Models for the core residues of Ce3 and Ce4 were created (B. A. W. *et al.*, manuscript in preparation) based upon homologous residues from IgG Mab231 (ref. 46). A new 2F_o–F_c map using phases from the receptor and core residues of Ce3 showed density for the two Ce4 domains. A model for the core residues in Ce4 was inserted. Rigid body refinement of the receptor, the core residues in Ce3, and the core residues in Ce4 produced an R_{free} of 45%. 2F_o–F_c maps and composite omit maps revealed electron density for protein and carbohydrate atoms absent from the model.

Refinement was continued with the form II data using all data with $|F| > 0$ up to the resolution shell where the $\langle I/\sigma \rangle$ dropped to 1.5 (3.25 Å). The data beyond 3.5 Å are weak, and we report the 3.5 Å structure and statistics here. Model building was done with the program O⁴⁷ and refinement carried out with CNS⁴⁸. The R_{free} set comprises 5% of the data that were selected using thin resolution shells. Strong non-crystallographic symmetry restraints (300 kcal mol⁻¹ Å⁻² on positions and 2 Å² on B-factors) were separately imposed on the Ce3 and Ce4 domains, while excluding the receptor-binding loops and the Ce3 carbohydrate. The ratio of the number of reflections to the number of atoms in the refinement is relatively high for this resolution (~5:1). A bulk solvent correction was applied and individual B-factors were used, as they lowered the R_{free} from 29.3% to 27.3% relative to grouped B-factors. The high overall B-factors are consistent with the weak diffraction from these crystals, but good electron density is observed throughout the model. After inserting all missing loops from the protein, five CHAPS molecules were located as large peaks of positive density in F_o–F_c maps. Of the 620 amino acids in the complex, 14 are missing from the model, all at chain termini: FcεRIα residues 174–176; Fc chain 1 residues 326, 327 and 545–547; and Fc chain 2 residues 326–328 and 545–547. Current refinement statistics are summarized in Table 1. Figures were made using the programs Molscript⁴⁹ and Grasp⁵⁰.

The form I crystal structure was solved by molecular replacement with AMoRe using the form II model of the complex. A single solution had a correlation coefficient of 26.0 with the next highest peak 13.5 (using normalized F_s). Searches in the enantiomorphic space group gave correlations lower than 12.3. Given the low resolution of form I crystals, refinement was limited to rigid body minimization.

Received 7 April; accepted 22 June 2000.

- Metzger, H. The receptor with high affinity for IgE. *Immunol. Rev.* **125**, 37–48 (1992).
- Kinet, J. P. The high-affinity IgE receptor (FcεRI): from physiology to pathology. *Annu. Rev. Immunol.* **17**, 931–972 (1999).
- Basu, M. *et al.* Purification and characterization of human recombinant IgE–Fc fragments that bind to the human high affinity IgE receptor. *J. Biol. Chem.* **268**, 13118–13127 (1993).
- Keown, M. B. *et al.* Hydrodynamic studies of a complex between the Fc fragment of human IgE and a soluble fragment of the FcεRI α-chain. *Proc. Natl Acad. Sci. USA* **92**, 1841–1845 (1995).
- Keown, M. B., Ghirlando, R., Mackay, G. A., Sutton, B. J. & Gould, H. J. Basis of the 1:1 stoichiometry of the high affinity receptor FcεRI–IgE complex. *Eur. Biophys. J.* **25**, 471–476 (1997).
- Daeron, M. Fc receptor biology. *Annu. Rev. Immunol.* **15**, 203–234 (1997).
- Turner, H. & Kinet, J. P. Signalling through the high-affinity IgE receptor FcεRI. *Nature* **402** (Suppl. B), 24–30 (1999).
- Ravetch, J. V. & Clynes, R. A. Divergent roles for Fc receptors and complement *in vivo*. *Annu. Rev. Immunol.* **16**, 421–432 (1998).
- Cookson, W. The alliance of genes and environment in asthma and allergy. *Nature* **402** (suppl. B), 5–11 (1999).
- Jardieu, P. M. & Fick, R. B. Jr IgE inhibition as a therapy for allergic disease. *Int. Arch. Allergy Immunol.* **118**, 112–115 (1999).
- Chang, T. W. The pharmacological basis of anti-IgE therapy. *Nature Biotechnol.* **18**, 157–162 (2000).
- Garman, S. C., Kinet, J. P. & Jardetzky, T. S. Crystal structure of the human high-affinity IgE receptor. *Cell* **95**, 951–961 (1998).
- Nissim, A., Jouvin, M. H. & Eshhar, Z. Mapping of the high affinity Fcε-receptor binding site to the third constant region domain of IgE. *EMBO J.* **10**, 101–107 (1991).
- Keegan, A. D., Fratazzi, C., Shopes, B., Baird, B. & Conrad, D. H. Characterization of new rat anti-mouse IgE monoclonals and their use along with chimeric IgE to further define the site that interacts with FcεRII and FcεRI. *Mol. Immunol.* **28**, 1149–1154 (1991).
- Presta, L. *et al.* The binding site on human immunoglobulin E for its high affinity receptor. *J. Biol. Chem.* **269**, 26368–26373 (1994).
- Henry, A. J. *et al.* Participation of the N-terminal region of Ce3 in the binding of human IgE to its high-affinity receptor FcεRI. *Biochemistry* **36**, 15568–15578 (1997).
- Granato, D. A. & Neeser, J. R. Effect of trimming inhibitors on the secretion and biological activity of a murine IgE monoclonal antibody. *Mol. Immunol.* **24**, 849–855 (1987).
- Letourneur, O., Sechi, S., Willette-Brown, J., Robertson, M. W. & Kinet, J. P. Glycosylation of human truncated FcεRI α-chain is necessary for efficient folding in the endoplasmic reticulum. *J. Biol. Chem.* **270**, 8249–8256 (1995).
- Kanellopoulos, J. M., Liu, T. Y., Poy, G. & Metzger, H. Composition and subunit structure of the cell receptor for immunoglobulin E. *J. Biol. Chem.* **255**, 9060–9066 (1980).
- Robertson, M. W. Phage and Escherichia coli expression of the human high affinity immunoglobulin E receptor α-subunit ectodomain. Domain localization of the IgE-binding site. *J. Biol. Chem.* **268**, 12736–12743 (1993).
- Maxwell, K. F. *et al.* Crystal structure of the human leukocyte Fc receptor, FcγRIIa. *Nature Struct. Biol.*

- 6, 437–442 (1999).
22. Sondermann, P., Huber, R. & Jacob, U. Crystal structure of the soluble form of the human FcγRIIb: a new member of the immunoglobulin superfamily at 1.7 Å resolution. *EMBO J.* **18**, 1095–1103 (1999).
23. Sondermann, P., Jacob, U., Kutscher, C. & Frey, J. Characterization and crystallization of soluble human FcγRII (CD32) isoforms produced in insect cells. *Biochemistry* **38**, 8469–8477 (1999).
24. Sanchez, L. M., Penny, D. M. & Bjorkman, P. J. Stoichiometry of the interaction between the major histocompatibility complex-related Fc receptor and its Fc ligand. *Biochemistry* **38**, 9471–9476 (1999).
25. Shi, J. *et al.* Interaction of the low-affinity receptor CD23/FcεRII lectin domain with the Fcε3-4 fragment of human immunoglobulin E. *Biochemistry* **36**, 2112–2122 (1997).
26. Deisenhofer, J., Colman, P. M., Epp, O. & Huber, R. Crystallographic structural studies of a human Fc fragment. II. A complete model based on a Fourier map at 3.5 Å resolution. *Hoppe Seylers Z. Physiol. Chem.* **357**, 1421–1434 (1976).
27. Harris, L. J., Larson, S. B. & McPherson, A. Comparison of intact antibody structures and the implications for effector function. *Adv. Immunol.* **72**, 191–208 (1999).
28. Schreiber, S. L. Target-oriented and diversity-oriented organic synthesis in drug discovery. *Science* **287**, 1964–1969 (2000).
29. Mallamaci, M. A. *et al.* Identification of sites on the human FcεRI α-subunit which are involved in binding human and rat IgE. *J. Biol. Chem.* **268**, 22076–22083 (1993).
30. Hulett, M. D., Witort, E., Brinkworth, R. I., McKenzie, I. F. & Hogarth, P. M. Identification of the IgG binding site of the human low affinity receptor for IgG FcγRII. Enhancement and ablation of binding by site-directed mutagenesis. *J. Biol. Chem.* **269**, 15287–15293 (1994).
31. Hulett, M. D., Witort, E., Brinkworth, R. I., McKenzie, I. F. & Hogarth, P. M. Multiple regions of human FcγRII (CD32) contribute to the binding of IgG. *J. Biol. Chem.* **270**, 21188–21194 (1995).
32. Cook, J. P. *et al.* Identification of contact residues in the IgE binding site of human FcεRIα. *Biochemistry* **36**, 15579–15588 (1997).
33. Hulett, M. D., Brinkworth, R. I., McKenzie, I. F. & Hogarth, P. M. Fine structure analysis of interaction of FcεRI with IgE. *J. Biol. Chem.* **274**, 13345–13352 (1999).
34. Sayers, I. *et al.* Amino acid residues that influence FcεRI-mediated effector functions of human immunoglobulin E. *Biochemistry* **37**, 16152–16164 (1998).
35. Anderson, C. L., Ryan, D. H., Looney, R. J. & Leary, P. C. Structural polymorphism of the human monocyte 40 kilodalton Fc receptor for IgG. *J. Immunol.* **138**, 2254–2256 (1987).
36. Duncan, A. R., Woof, J. M., Partridge, L. J., Burton, D. R. & Winter, G. Localization of the binding site for the human high-affinity Fc receptor on IgG. *Nature* **332**, 563–564 (1988).
37. Chappel, M. S. *et al.* Identification of the Fcγ-receptor class I binding site in human IgG through the use of recombinant IgG1/IgG2 hybrid and point-mutated antibodies. *Proc. Natl Acad. Sci. USA* **88**, 9036–9040 (1991).
38. Hulett, M. D., McKenzie, I. F. & Hogarth, P. M. Chimeric Fc receptors identify immunoglobulin-binding regions in human FcγRII and FcεRI. *Eur. J. Immunol.* **23**, 640–645 (1993).
39. Jefferis, R., Lund, J. & Pound, J. D. IgG-Fc-mediated effector functions: molecular definition of interaction sites for effector ligands and the role of glycosylation. *Immunol. Rev.* **163**, 59–76 (1998).
40. Keown, M. B., Henry, A. J., Ghirlando, R., Sutton, B. J. & Gould, H. J. Thermodynamics of the interaction of human immunoglobulin E with its high-affinity receptor FcεRI. *Biochemistry* **37**, 8863–8869 (1998).
41. Zheng, Y., Shopes, B., Holowka, D. & Baird, B. Dynamic conformations compared for IgE and IgG1 in solution and bound to receptors. *Biochemistry* **31**, 7446–7456 (1992).
42. Dorrington, K. J. & Bennich, H. H. Structure–function relationships in human immunoglobulin E. *Immunol. Rev.* **41**, 3–25 (1978).
43. Otwinowski, Z. & Minor, W. in *Methods in Enzymology: Macromolecular Crystallography, Part A* (eds Carter, C. W. Jr & Sweet, R. M.) 307–326 (Academic, New York, 1997).
44. Collaborative Computational Project No. 4. The CCP4 Suite: Programs for Protein Crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
45. Kissinger, C. R., Gehlhaar, D. K. & Fogel, D. B. Rapid automated molecular replacement by evolutionary search. *Acta Crystallogr. D* **55**, 484–491 (1999).
46. Harris, L. J., Larson, S. B., Hasel, K. W. & McPherson, A. Refined structure of an intact IgG2a monoclonal antibody. *Biochemistry* **36**, 1581–1597 (1997).
47. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
48. Brunger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
49. Kraulis, P. J. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
50. Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct. Funct. Genet.* **11**, 281–296 (1991).

Acknowledgements

We thank K. Baker, J. Quintana, and D. Keane for help with synchrotron data collection and Y. Wang, G. Sun, and E. Song for technical assistance. Portions of this work were performed at the Dupont-Northwestern-Dow (DND-CAT) Synchrotron Research Center of the Advanced Photon Source and the Advanced Light Source beamline 5.0.2. This research has been supported by the National Institutes of Health (T.S.J.), the Pew Scholars Program in the Biomedical Sciences (T.S.J.), Heska Corporation (T.S.J.), and the American Cancer Society (S.C.G.). Coordinates have been deposited in the PDB with accession number 1F6A.

Correspondence and requests for materials should be addressed to T.S.J. (e-mail: tedj@northwestern.edu).

The 3.2-Å crystal structure of the human IgG1 Fc fragment–Fc γ RIII complex

Peter Sondermann*, Robert Huber*, Vaughan Oosthuizen*† & Uwe Jacob*

* Max-Planck-Institut für Biochemie, Abteilung Strukturforschung, Am Klopferspitz 18a, D-82152 Martinsried, Germany

† University of Port Elizabeth, Department of Biochemistry & Microbiology, PO Box 1600, Port Elizabeth 6000, South Africa

The immune response depends on the binding of opsonized antigens to cellular Fc receptors and the subsequent initiation of various cellular effector functions of the immune system. Here we describe the crystal structures of a soluble Fc γ receptor (sFc γ RIII, CD16), an Fc fragment from human IgG1 (hFc1) and their complex. In the 1:1 complex the receptor binds to the two halves of the Fc fragment in contact with residues of the C γ 2 domains and the hinge region. Upon complex formation the angle between the two sFc γ RIII domains increases significantly and the Fc fragment opens asymmetrically. The high degree of amino acid conservation between sFc γ RIII and other Fc receptors, and similarly between hFc1 and related immunoglobulins, suggest similar structures and modes of association. Thus the described structure is a model for immune complex recognition and helps to explain the vastly differing affinities of other Fc γ R–IgG complexes and the Fc ϵ R1 α –IgE complex.

Several important functions of the immune system are associated with the Fc portion of immunoglobulin molecules and are assigned to different regions of this structurally well characterized protein¹. These functions are associated with binding to different proteins like complement, the immunologically active Fc receptors (Fc γ Rs), the neonatal Fc receptor (FcRn) which transports maternal IgG to the fetus, and the bacterial Fc receptors, protein A and protein G, which are believed to mask the bacteria through immobilized immunoglobulins on their surface.

The main immunological activity of the human Fc (hFc) is exerted through binding to the membrane bound Fc γ R. Once an immunogenic particle is opsonized by IgG it crosslinks Fc γ Rs on the cell surface and triggers the immune response². While monomeric Fc γ Rs are not able to activate the immune system, the intracellular parts of aggregated Fc γ Rs recruit different SH2-domain-containing molecules which may result in stimulation or repression of the immune response³. The cellular responses are complex and range from endocytosis and antibody-dependent cellular cytotoxicity (ADCC), to the secretion of inflammatory mediators, enhanced antigen presentation and regulation of the antibody production. Which of these are triggered depends on the particular Fc γ R isoform and the type of cell⁴.

Fc γ Rs are homologue type I membrane proteins that are expressed on all immunocompetent cells. The higher affinity of the Fc γ RI (CD64, $K_A \approx 10^8 \text{ M}^{-1}$) compared to the tandem immunoglobulin domain Fc γ Rs (Fc γ RII, CD32; Fc γ RIII, CD16; $K_A \approx 10^6$ – 10^7 M^{-1}) is attributed to its third domain. In addition to the transmembrane form of the Fc γ RIII, which is expressed on natural killer cells and monocytes (Fc γ RIIIa), a glycosylphosphatidylinositol-anchored

form is present on neutrophils (Fc γ RIIIb). Besides these membrane bound forms of Fc γ Rs, soluble receptors exist *in vivo*, being produced by alternative splicing, separate genes or proteolysis⁵ but their exact functions remain unknown.

Fc γ Rs have a key role in the immune system by connecting the humoral and cellular immune response. Therefore, binding of IgG to Fc γ R represents a crucial mechanism in immunology, but it is not well understood on the molecular level. The structures of the Fc fragment in complex with protein A⁶, protein G⁷ and rat FcRn⁸ revealed binding sites on the C γ 2/C γ 3 interface. Mutagenesis and domain swapping studies did not identify the binding area of Fc and Fc γ R conclusively and, besides binding sites in the lower hinge region^{9–12}, contacts of the C γ 2/C γ 3 interface with the Fc γ R^{13,14} were proposed. The structure of the soluble extracellular part of a Fc γ R¹⁵ and the wealth of biochemical functional data (for a review see ref. 16) do not allow prediction of the complex structure with confidence as extensive domain rearrangements during complex formation must be considered. Modelling studies on the basis of the Fc γ R and Fc fragment structures indicate that either the C γ 2/C γ 3 interface¹⁵ or the lower hinge region¹⁷ are the principal site of interaction. The structure of a Fc γ R–Fc fragment complex presented here provides a sound basis for the modelling of other FcR–Fc complexes and their functional interpretation.

The overall structure

The crystal structure of the genetically engineered, membrane-anchor-free sFc γ RIII in a 1:1 complex with the monoclonal human Fc fragment of IgG1 (hFc1) was solved by molecular replacement using the previously analysed structures of the complex

Table 1 Summary of crystallographic data

Protein	sFc γ RIII	hFc1	sFc γ RIII–hFc1
Recording temperature	291 K	100 K	100 K
Spacegroup	P2 ₁ (1)2(1)	P2(1)2(1)2(1)	P6(5)22
Cell constants	$a = 36.70 \text{ \AA}$, $b = 60.29 \text{ \AA}$, $c = 85.58 \text{ \AA}$, $\alpha = \beta = \gamma = 90.0^\circ$	$a = 49.57 \text{ \AA}$, $b = 78.80 \text{ \AA}$, $c = 140.40 \text{ \AA}$, $\alpha = \beta = \gamma = 90.0^\circ$	$a = b = 115.32 \text{ \AA}$, $c = 299.10 \text{ \AA}$, $\alpha = \beta = 90.0^\circ$, $\gamma = 120.0^\circ$
Molecules per asymmetric unit/solvent content	1/48%	1/49%	1/67%
No. of independent reflections/multiplicity	6725/3.0	13290/3.4	18596/3.4
Completeness overall/last shell	94.7%/72.3%	95.0%/74.5%	94.9%/66.8%
R_{merge} overall / last shell	11.4%/50.0%	12.0%/52.7%	13.3%/60.7%
Limiting resolution (2 I/σ)	2.5 $\text{\AA}$	2.7 $\text{\AA}$	3.2 $\text{\AA}$
$R_{\text{factor}}/R_{\text{free}}^*$	19.5/26.1	27.0/35.3	27.08/35.69
RMSD bonds/RMSD angles	0.012/1.64	0.009/1.88	0.016/2.173
Overall B factor ($\text{\AA}^2$)	29.7	47.6 (21.3, 1FC1 at 291 K)	95.1

* R_{free} : 5% of the reflections were used as a reference data set and were not included in the refinement.

components (Table 1). The structures of hFc1 and polyclonal hFc^{1,6} are almost identical. The structure of sFcγRIII shows the characteristic heart-shaped domain arrangement described for sFcγRIIb¹⁵ and other sFcRs^{18,19}.

The protein cores without carbohydrates were used to generate phases and the calculated electron density maps produced clearly traceable difference density at the expected position of the sugar moiety and the hinge region, confirming the analysis.

In the complex structure, the receptor is bound between the two Cγ2 domains of the Fc fragment. Only domain 2 of the sFcγRIII and two residues from the linker connecting this domain with domain 1 interact in the complex with different regions of both Cγ2 domains

(Cγ2-A and Cγ2-B) and the preceding hinge regions of hFc1 (Fig. 1). The buried surface upon complex formation is 350 Å² for the Cγ2-A and 545 Å² for the Cγ2-B domain. Both complex components undergo a substantial domain rearrangement upon binding. Structural rearrangements within the domains are restricted to the loops involved in complex formation. The sFcγRIII molecule experiences an opening of its interdomain angle from 70° to 80° and the solvent exposed side chain Trp 95 rotates into the opened cleft (Fig. 2a) and contacts Tyr 14.

The horseshoe shaped hFc1 molecule also opens (Fig. 2b). The Cγ2 domains bend away from the C2-axis of the homodimeric hFc1 and enlarge the distance between the Pro 329 residues located on the

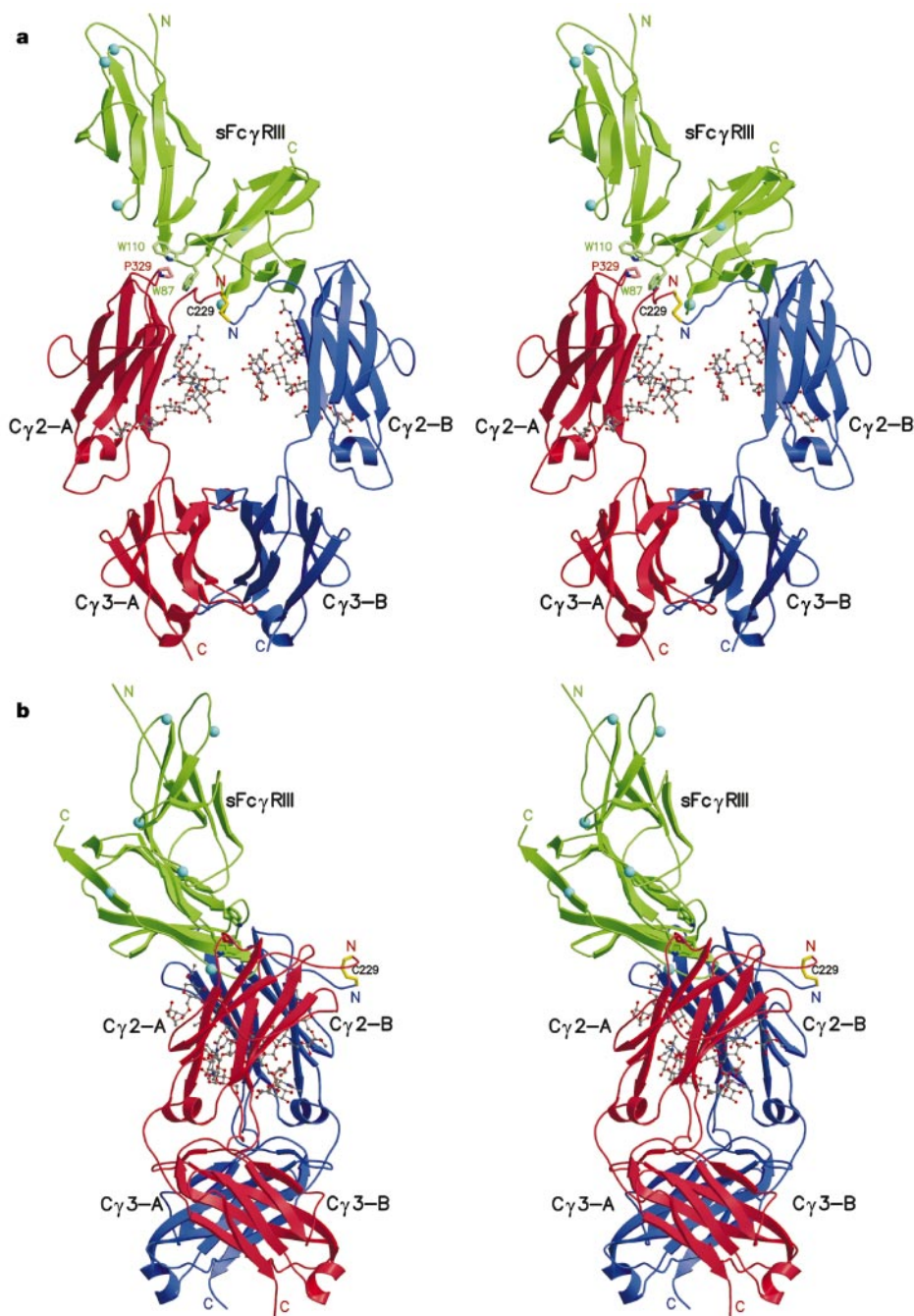


Figure 1 The overall structure of the sFcγRIII-hFc1 complex. **a**, Stereo ribbon representation with the dimer axis of hFc1 (red and blue) orientated vertically. The 'proline sandwich' consisting of Pro 329 of the Cγ2-A domain and Trp 87 and Trp 110 of sFcγRIII (green) is shown in ball and stick together with the carbohydrate residues of the Fc

fragment and the interchain disulphide bridge (yellow) of the Cys 229 residues. The potential *N*-glycosylation sites in sFcγRIII are shown as cyan balls. **b**, Side view of the complex obtained by a rotation of 90° around the *y* axis. Unless indicated, all figures were produced with the programs MOLSCRIPT⁴⁶ and RASTER3D⁴⁷.

amino-terminal tip of both C γ 2 domains by 7 Å, and the orientation of the C γ 3 domains remains unchanged. The C γ 2-B domain is significantly more dislocated upon binding of sFc γ RIII than the C γ 2-A domain. The introduced asymmetry is reflected in the stoichiometry of complex formation (see below).

In the uncomplexed hFc1, the complete N-linked carbohydrate moieties attached at Asn 297 of both chains can be traced in the electron density. They contact each other in the central cavity of hFc1 holding the C γ 2 domains at a minimum distance. Although this interchain carbohydrate contact is disrupted upon sFc γ RIII binding, all sugar residues remain traceable in the electron density, indicating the role of structural stabilization by carbohydrate–protein interactions. The distance between the O4 atoms of the α -D-mannose (MAN)8 residues increases from 2.8 Å to 5.3 Å upon complex formation.

The contact interface between sFc γ RIII and hFc1

sFc γ RIII binds hFc1 with the B/C loop (Trp 110 to Ala 114) and F/G loop (Val 155 to Lys 158) and with the C strand (His 116 to Thr 119) and C' strand (Asp 126 to His 132) on its carboxy-terminal domain 2. Additionally, Arg 152 and the connector between the N-terminal domain 1 and domain 2 (Ile 85 to Trp 87) is involved in binding (Fig. 3). These regions interact with loops on the hFc1 within C γ 2-B (B/C: Asp 265 to Glu 269, C'/E: Asn 297 to Thr 299), with the F/G loop of C γ 2A (Ala 327 to Ile 332), with the carbohydrate residue N-acetyl-D-glucosamine (NAG)1 of C γ 2-B and with the hinge region of both heavy chains (Leu 234 to Ser 239).

The interactions consist mainly of van der Waals contacts and six potential hydrogen bonds most of which are contributed by four residues of the C γ 2-B hinge (Leu 235 to Ser 239) emphasising the importance of this region (Fig. 4). Two main contact areas exist. First, Pro 329 of C γ 2-A is engaged tightly by Trp 87 and Trp 110 of the receptor ('proline sandwich', Figs 1 and 4). Second, residues Leu 234 to Ser 239 of the lower hinge of both hFc1 chains, that were already recognized as important for Fc γ R binding by mutagenesis studies^{9,20}, are engaged in sFc γ RIII binding. These residues were disordered in the hFc1 crystal structures but are defined in the complex.

The C γ 2-domains of both chains follow a 180° rotation around the γ axis (Fig. 1) except for residues preceding Pro 238 which run perpendicular to the hFc1 axis. Thus, the hinge peptides of both chains are found in differing conformations in the complex. Furthermore, the receptor and the Fc fragment are in very close

contact at this position and glycines are required for steric and conformational reasons (Fig. 4). Gly 237 of C γ 2-A and Gly 236 of C γ 2-B adopt Phi/Psi angles in the complex not allowed for other residues. The hinge residues Leu 234 to Ser 239 of C γ 2-A are in contact with residues Thr 113, Ala 114 and Val 155 to Lys 158 of the receptor. The main contact is mediated by Leu 235 which binds in a shallow hydrophobic groove with Val 155 at its bottom and the Ala 114 CB and Thr 113 CG2 at its rim.

The hinge residues of the C γ 2-B domain are bound in a narrow channel of sFc γ RIII lined by His 116 and His 132 on one side and Lys 117 on the other. The glycines 236 and 237 of the hinge peptide bind into this channel and Tyr 129 and His 131 form additional contacts to the hinge peptide. In this arrangement the histidine residues 116, 131 and 132 are potential hydrogen bond partners to hinge residues (Fig. 4). Leu 235 is in contact with His 116 and His 132 and its importance for binding is not reflected in this binding mode, but in other receptors this region is considerably more hydrophobic, as discussed below.

The receptor contacts the interstrand loops of the C γ 2-B domain and the terminal residues of the β -strands. Here the side chains of the amino acid residues Asp 126 to His 131 (C' strand) are bound to the amino acid residues Asp 265 to Glu 269 and Asn 297 to Thr 299 of C γ 2-B. One more potential hydrogen bond is found in this area formed by Arg 152 located on the F strand to the carbohydrate residue NAG1 which is directly attached to Asn 297 of C γ 2-B. Some more distant contacts to this sugar residue are formed by Lys 117, Thr 119, Asp 126 and Tyr 129. The influence of the Fc glycosylation upon Fc γ RIII binding was shown with deglycosylated IgG which still binds Fc γ RIII, albeit with a reduced affinity²¹.

Our sFc γ RIII preparation is unglycosylated. One potential N-glycosylation site at Asn 159 is located close to the binding region (Figs 1 and 4). The Asn side chain points away from the binding site but a larger carbohydrate moiety in the sFc γ RIII may influence the affinity to IgG. The affinity of the crystallized sFc γ RIII for hFc1 ($K_A = 5 \times 10^5 \text{ M}^{-1}$; K. Maenaka, personal communication) is similar to sFc γ RIIB²² and in the range for glycosylated sFc γ RIII²³. However, the affinity of Fc γ RIII to IgG *in vivo* seems to be dependent on the kind of receptor glycosylation²⁴.

The stoichiometry of the sFc γ RIII-hFc1 complex

In the crystal, we observed a 1:1 stoichiometric complex of the sFc γ RIII and hFc1 components. Binding of a second sFc γ RIII molecule in a related manner on the opposite side of the Fc fragment

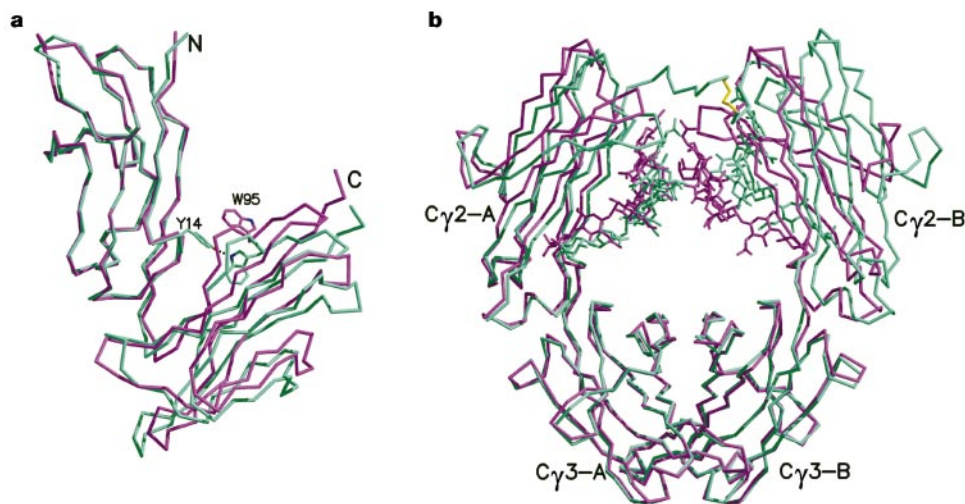


Figure 2 Superposition of the complexed components with the free structures (C α trace). **a**, sFc γ RIII in complexed form (cyan) was superimposed with a least square algorithm using the C α atoms of the first domain onto the structure of sFc γ RIII obtained from sFc γ RIII crystals (magenta). Trp 95 and Tyr 14 which form a new hydrogen bond upon

complex formation are shown in ball and stick representation **b**, Overlay of the hFc1 structures obtained from the complex (cyan) and from hFc1 crystals (magenta). For the superposition the C α atoms of both C γ 3 domains were used. The orientation of the structures is the same as in Fig. 1a.

would interfere with crystal packing, but is not excluded for steric reasons in the isolated molecule. The perspective in Fig. 1b indicates sufficient space for such a binding. However, the asymmetry of the Fc fragment enforced by the binding of one FcγRIII molecule would disallow the same binding mode for a second receptor molecule and possibly prevent formation of a 2:1 complex. A 1:1 sFcγRIII-hIgG complex was found in solution by ultracentrifugation experiments²³. Additionally, the crosslinking of two FcRs on the membrane by one antibody can already trigger an immune response, as in the case of IgE²⁵. A 1:1 stoichiometry avoids a permanent stimulation of the immune system by the monomeric immunoglobulins present at high concentrations in the serum.

The sFcγRIII-hFc complex: A model for FcR-Ig complexes

A sequence comparison of sFcγRIII with other FcγRs and the FcεRIα shows a high degree of similarity (Fig. 3). Mutagenesis studies with FcγRII^{18,26} and FcεRIα^{27,28} indicate that the same regions are involved in complex formation as are defined here for sFcγRIII. Similarly, the sequences of the four human and mouse IgG subclasses are almost identical, while the homology between IgE and IgG is above 50%. The most diverse regions are the hinge residues N-terminal to the Cγ2 domain. Together with interstrand loops of the Cγ2 domains these regions are essential for sFcγRIII binding. Within these short regions, conserved interactions as well as non-conserved contacts are possible, which might be important in specificity generation.

On the basis of the conserved sequences within the receptor and the different IgGs and IgE, respectively, we propose a common model describing the principal interactions within the various complexes. The proline sandwich appears as the primary binding motive in all FcR-Ig complexes. The two tryptophan residues are

conserved in the FcRs including FcεRIα and the proline occurs in all IgGs and IgE. Further conserved contacts are Lys 117 to Asp 265, Thr 119 to NAG1 and Lys 128 to Glu 269 (the numbers in all IgGs and FcRs refer to the homologous position in hIgG1 and sFcγRIII, respectively, Fig. 3). The hinge peptide is a second central recognition site. The sequence variation in this region is likely to be the main reason for the differing affinities in the FcR-Ig pairs. The interactions seem to be modulated in different receptor Ig pairings in a productive or destructive way by the non-conserved regions of the binding loops.

The FcγRII interaction with IgG was studied by NMR with multiply labelled hFc²⁹, and the same contact areas were identified as in the crystalline sFcγRIII-hFc1 complex. FcγRII and FcγRIII are 50% identical and the differences affect the loops in contact with the hinge, but not the contact regions to Cγ2-A and Cγ2-B. The glycines 236 and 237 of the Cγ2-B domain hinge peptide are bound in the narrow channel between His116, His132 and Lys 117. Lys 117 is conserved in FcγRII but the histidine residues are replaced by Val and Leu, respectively. These exchanges create a small hydrophobic patch to which Leu 235 may bind and allow a tighter contact of receptor and hinge peptide. The FcγRIIa high responder allele (HR), named after the ability of T cells from some individuals to respond to murine IgG1 induced mitogenesis³⁰, has an arginine at position 131. This residue is a histidine in the FcγRIIa low responder (LR). The affinity of the FcγRIIa-HR for all human IgG subclasses is reduced compared with the LR allele³¹ and almost abolished for IgG2. A plausible explanation may be steric clashes between the larger side chain of Arg 131 in the receptor and Pro 238 of the hinge peptide, with associated structural rearrangements. A rationale for the higher affinity of the HR allele for mouse IgG1 is offered by the replacement of Pro 238 by the smaller serine which

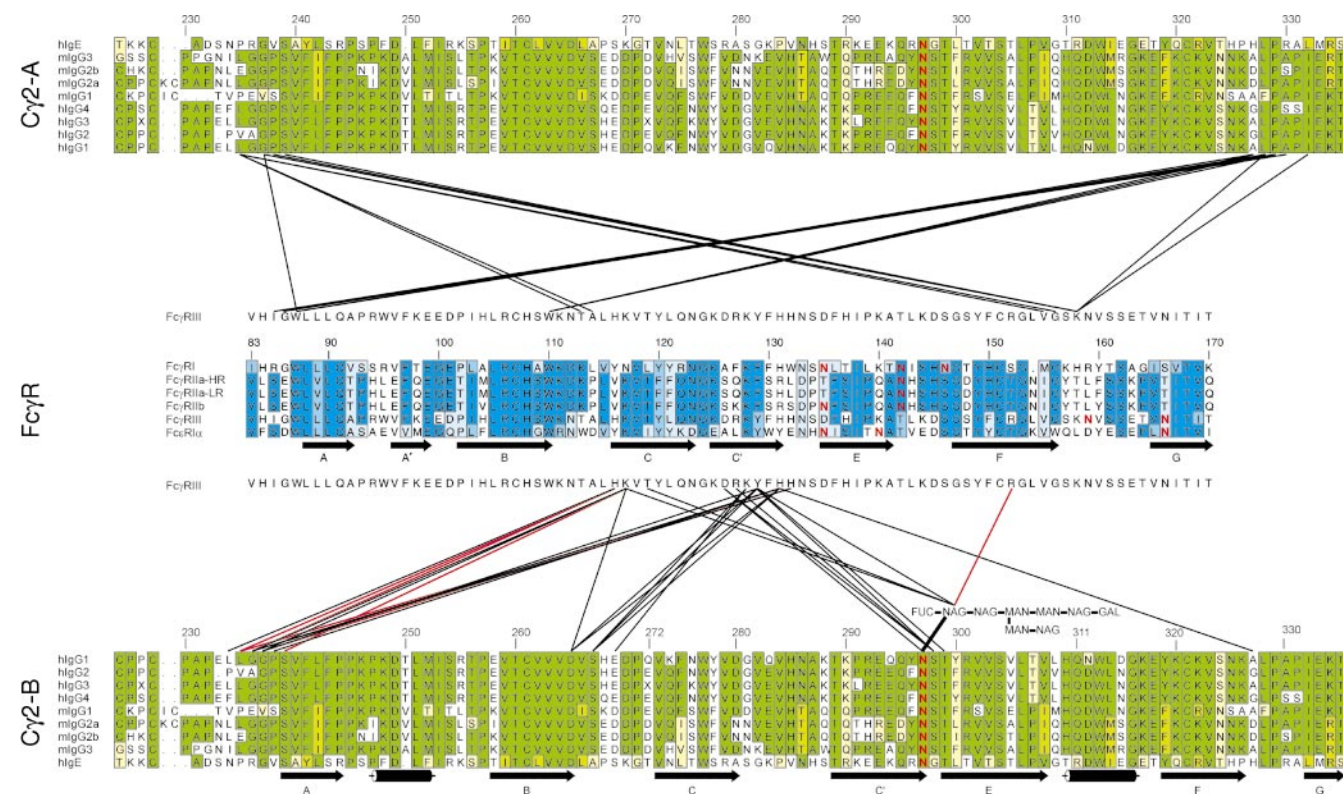


Figure 3 Structure based sequence alignment of the complex components. The regions involved in complex formation between sFcγRIII (middle) and both Cγ2 domains of hFc1 (top and bottom) are aligned to their homologue proteins, human FcRs and human and mouse immunoglobulins, respectively. The degree of similarity correlates with the depth of the colouring. Potential N-glycosylation sites are depicted in red letters and the secondary structure elements are drawn below the sequences. Amino acid residues in

van der Waals contacts between sFcγRIII and hFc1 are connected with black lines whereas hydrogen bonds are coloured in red. The contacts of sFcγRIII to the carbohydrate moiety attached to Asn 297 of Cγ2-B are also indicated in this alignment. The numbering is according to the complex constituent sequences (for example, FcγRI does not have a sequence position 154). The figure was created using the program ALSCRIPT⁴⁸. GAL, D-galactose; FUC, fucose.

could even form a hydrogen bond to Arg 131. Fc γ RIIa-LR binds hIgG2 possibly because the hinge contacting residues Val 116 and Leu 132 are smaller than the histidines in sFc γ RIII, and are able to accommodate Ala 236.

The significantly higher affinity of sFc γ RI is thought to be mediated by its third domain because the two N-terminal domains show an affinity to IgG comparable to that of Fc γ RII/III³². Both Fc γ RI N-terminal domains are closely related to the Fc γ RII/III domains and will probably bind the Fc-fragment, as in the sFc γ RIII-hFc1 complex. The hinge mutation E235L increases the affinity of mIgG2b by more than 100-fold⁹, which underlines the central importance of this residue for Fc γ RI binding. Modelling indicates that Leu 235 of both C γ 2 chains may be bound to an extended hydrophobic patch in Fc γ RI formed by Trp 132 and Tyr 116. The role of the extra domain of Fc γ RI in enhanced IgG binding remains unclear. It could contribute to the affinity by a stabilization of the open receptor conformation or through direct binding to the Fc fragment.

The Fc fragment of IgE consists of three constant domains; C ϵ 2, C ϵ 3 and C ϵ 4. The C ϵ 3 domain (homologue to the C γ 2 domain of IgG) has been proposed as the binding site for Fc ϵ RI α ³³. The stoichiometry of the sFc ϵ RI α -IgE complex is 1:1 and complex formation occurs without significant structural change in the domains of either component³⁴. We modelled the putative sFc ϵ RI α -IgE interaction by replacing all amino acids of the sFc γ RIII-hFc1 complex with their homologues. No steric clashes

were seen. The critical differences were again found in the hinge region. Here, Pro 238 is replaced by valine which may result in a displacement of the hinge. This is comparable to the Pro 238 Ser exchange in mouse IgG1, as for the HR/LR dualism of Fc γ RIIa. The extra space gained on the C ϵ 3-B domain may be filled by Tyr 131 of the receptor and Arg 236 of hIgE (Gly in hIgG1). Mutation of this arginine to serine results in a 120-fold decrease in affinity³⁵. In our model, Arg 236 of C ϵ 3-B would adopt an upward orientation and form two additional salt bridges to Asp 118 and Glu 136 of Fc ϵ RI α . The replacement of Leu 235 with Pro in IgE yields additional space that could be filled by Tyr 116 (comparable to Tyr 116 in Fc γ RI) and two additional tryptophan residues Trp 113 and Trp 156, thereby extending the interaction area of the conserved proline sandwich. Trp 156 would insert between the Pro 235 residues of both hinge peptides to compensate for the exchanged Leu. This tryptophan rich region formed by the four Trp residues 87, 110, 113 and 156, may contribute much of the energy of complex formation when buried.

Comparison with other FcRs

The sFc γ RIII-hFc1 complex structure clarifies that the binding sites for Fc γ Rs on IgG do not overlap with other Fc receptors like Protein A, Protein G and FcRn. The structural reorganisation of hFc1 upon sFc γ RIII binding results in a dislocation of the C γ 2 relative to the C γ 3 domains. The relative domain arrangement is different in crystals of the free Fc fragment and in Fc fragment complexes

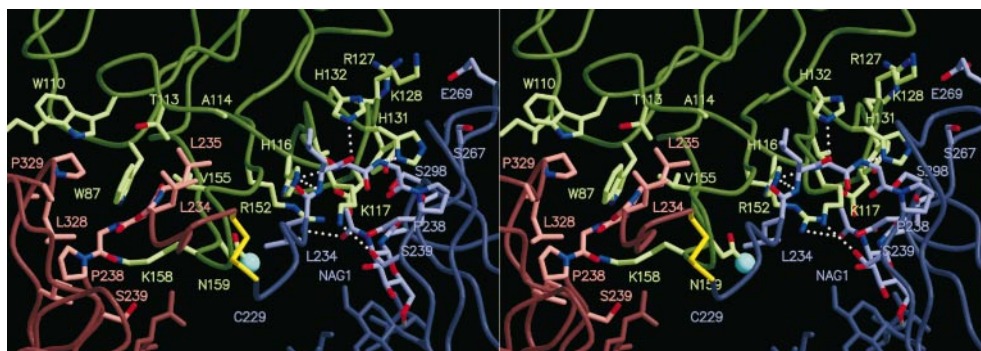


Figure 4 The binding site. A close up of the residues involved in the sFc γ RIII-hFc1 contact in stereo representation. The orientation of the complex is the same as in Fig. 1a. The backbone residues of sFc γ RIII, C γ 2-A and C γ 2-B are coloured green, red and blue,

respectively. Potential hydrogen bonds between sFc γ RIII and hFc1 are shown as dashed lines, potential N-glycosylation sites as cyan balls and the interdomain disulphide bridge between the Cys 229 of C γ 2-A and C γ 2-B in yellow.

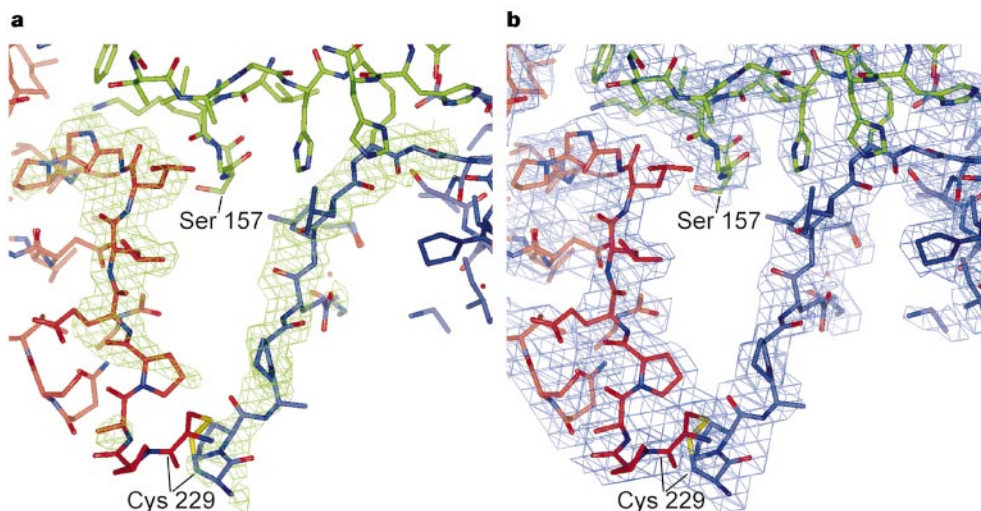


Figure 5 Omit map of the hinge region. **a**, Fo-Fc annealed omit map contoured at 2.3 σ covering the hinge region **b**, Final 2Fo-Fc electron density map contoured at 0.8 σ covering the final model. The colour coding and the orientation of the figure is the same as

in Figs 1 and 4 with the sFc γ RIII in green, the hFc1 C γ 2-A in red and the C γ 2-B domain in blue. Ser 157 of sFc γ RIII and both Cys 229 of hFc1 are labelled.

with protein A, protein G and FcRn. An angular change between Cy2 and Cy3 could therefore provide a mechanistic basis for an allosteric interaction between different receptors, such that antibodies bound through FcRn could not bind FcγRs or vice versa. Whether this effect exists or has physiological relevance, needs to be clarified.

Conclusions

The FcR–Ig interaction is a promising target for the therapy of allergies, autoimmune diseases or graft versus host rejections. Structure-based drug design might help to find small molecules that inhibit the FcR–Ig interaction and downregulate the immune system. Substances interfering with the proline sandwich could downregulate the entire immune system while inhibiting the interactions in the hinge region and could yield a receptor-specific block. Specific inhibition of the FcεRIα–IgE interaction, for instance, might help for the treatment of allergies. Also a specific inhibition of FcγRIII in the case of autoimmune diseases³⁶ or FcγRIIb in tumour therapies³⁷ could provide new strategies for the treatment of these diseases and for the understanding of the role of these receptors in the immune system.

The engineering of therapeutic antibodies of low or high receptor affinity that block cellular proteins or target the antigen to an FcR subclass seems feasible. Monoclonal antibodies with a reduced affinity to FcγR which downregulate T-cell activation, are already in clinical trials, such as an anti-CD3 antibody for the treatment of acute renal allograft rejection³⁸ and an anti-CD4 antibody for the treatment of autoimmune diseases³⁹. Intact Fc parts are indispensable for a long half-life of the antibodies *in vivo*. On the other hand the immune system should not be activated through FcγR binding, which is accomplished by the exchange of the [Leu 234]-Leu-Gly-[Gly 237] motif for Ala-Ala-Gly-Gly in the first example and for Phe-Glu-Gly-Gly in the latter.

It could also be advantageous to label antigens with antibodies of artificially enhanced affinity to FcγRI to trigger an effective immune response. So far this could only be realised by bispecific antibodies that recognize a tumour marker with one Fab and FcγRI with the other⁴⁰.

In this report, we showed that characteristic interaction sites like the proline sandwich are likely to be present in all FcRs and their respective Ig ligands. The knowledge of the Fc–FcR interactions facilitates the development of immune modulating compounds. □

Methods

Purification of the complex components and their crystallization

We produced sFcγRIII, comprising amino acids –4 to 171 of the mature protein (H₂N–MRT...TQG–COOH, in the sFcγRII numbering system, Fig. 3), by refolding of inclusion body protein produced in *Escherichia coli* as described⁴¹. We isolated the hIgG1 from the plasma of a myeloma patient by standard methods. The hIgG1 Fc fragment was prepared by plasmin digestion and purified by chromatography on protein A agarose and subsequent gel filtration and shows a clean N-terminus starting with amino acid 223 (H₂N–THTCPPC..., the amino acid sequence numbering is based on the Eu amino acid sequence⁴²) proven by N-terminal sequencing. For crystallization trials we mixed equimolar amounts of the proteins and adjusted their concentration to 10 mg ml^{–1} in 2 mM MOPS, 150 mM NaCl, 0.02% sodium azide, pH 7.1. We obtained the initial hexagonal crystals by an extended factorial screen in sitting drops after several days, using a pipetting robot (Crybot 3000, Qiagen, Hilden/Germany). The crystals used for data collection were grown in sitting drops using 2 μl of protein and 1 μl of reservoir solution (100 mM MES/Tris, pH 5.6, 6% PEG 8000, 200 mM K/Na tartrate, 0.02% sodium azide at 18 °C) and grew to their final size of 200 × 200 × 400 μm³ within one week.

Diffraction data collection

Crystals of space group P6₅22 with one complex molecule (1:1 stoichiometry; sFcγRIII:hFc1) in the asymmetric unit diffracted to 6 Å and decayed rapidly in the X-ray beam at 291 K. However, crystals diffracted to 3.0 Å when slowly accustomed to crystallization buffer containing 30% glycerol and flash cooled to 100 K. A complete data set was recorded at beamline BW6 at DESY Hamburg/Germany employing a Mar Research CCD detector and a wavelength of 1.050 Å. During the cryotreatment the cell constants of the crystals transform from $a = b = 114.06$ Å, $c = 303.77$ Å to $a = b = 115.32$ Å, $c = 299.10$ Å.

The resolution of the diffraction pattern increases significantly by this transformation but some disorder is also introduced. The images showed a strong anisotropic diffraction pattern and rather high R_{merge} , especially in higher resolution shells (Table 1) as seen in the subsequent data reduction (CCP4 program package⁴³).

Structure determination

All structures were solved with the methods of molecular replacement using the program AMoRe⁴⁴. The structure of the sFcγRIII–hFc1 complex was solved with the crystal structures of the isolated hFc1 (2.7 Å) and the sFcγRIII (2.5 Å) as search models (P.S., R.H. and U.J., unpublished data). Owing to the large structural reorganisation of the molecules upon complex formation, AMoRe yielded a single solution in space group P6₅22 for the positioned Fc fragment only when the resolution was lowered to 6.5 Å (Correlations coefficient = 40.1; next highest 34.0). With the positioned Fc part, the receptor was found with a clear solution (Correlations coefficient = 49.7, next highest 42.4). (Comparable results were obtained when the PDB entries 1FC1, human Fc fragment of IgG, and 2FCB, soluble human FcγRIIb, were employed as search models.) The positioned molecules fitted into the asymmetric unit without steric clash.

By a refinement considering each domain as rigid unit the analysis was extended stepwise to the resolution of 3.2 Å and the large relative movements of the individual domains became apparent. An overall B -factor refinement resulted in high B -factors with a large anisotropic contribution, in accordance with the prediction from the Wilson plot, indicating disorder caused by the cryo-treatment. This might be an intrinsic property of the Fc fragment, which has also been observed in other complex structures where parts of the Fc fragment were found disordered. Furthermore, cryo-cooled hFc1 crystals showed a similar effect, the overall B -factor increased to 48 Å² compared to 21 Å² for room temperature data (Table 1). Overall B -factors, assigned individually to the Cy2-A/Cy2-B (B -factor = 105/61 Å²), Cy3-A/Cy3-B (B -factor = 115/113 Å²) domains of hFc1 and sFcγRIII (B -factor = 83 Å²) in the complex structure, showed that significantly more disorder is introduced by the Cy3 domains. Thus, non-crystallographic symmetry (NCS) restraints were assigned to four groups, to the main chain and side chain atoms of both Cy2 domains and to the main chain and side chain atoms of both Cy3 domains. The refinement was continued employing simulated annealing and structure factor averaging routines from the CNS program package⁴⁵. Finally, individual B -factors were refined with strong restraints for bonded atoms. In the refinement the weights for the crystallographic term, for the NCS restraints of atomic positions and of B -values, and for the restraint of B -values of bonded atoms were chosen such that R_{free} runs through a shallow minimum. Because of these difficulties the refinement converged at rather high R -factors. Nevertheless, the calculated electron density maps were of surprising quality and clear difference density for the carbohydrate moiety and parts of the hinge region which have not been used for phasing became visible (Fig. 5). The calculated electron density at 3.2 Å was of sufficient quality to trace the backbone of the polypeptide chain and to rebuild amino acid side chains. However, some ambiguities in the main chain peptide orientations remained and questionable dihedral angles were left in the conformation of the search models. The dihedrals of all amino acid residues refined to allowed values in the Ramachandran plot except for Ala 231 of Cy2-A and Lys 4 and Glu 37 of sFcγRIII which were built into weak electron density. The model was completed in alternating cycles of model building and refinement.

The final model consists of the amino acid residues Asp 1 to Gly 172 of the sFcγRIII and Cys 229 to Ser 444 of both hFc1 chains. The segment Cys 229 to Glu 233 of both Cy2 chains was built into weak electron density (Fig. 5) while the terminal residues Thr 223 to Pro 228 and Pro 445 to Gly 446 could not be traced at all. In the sFcγRIII the N-terminal residues Met –4 to Glu –1 were also not visible in the electron density.

Received 17 March; accepted 1 June 2000.

- Huber, R., Deisenhofer, J., Colman, P. M., Matsushima, M. & Palm, W. Crystallographic structure studies of an IgG molecule and an Fc fragment. *Nature* **264**, 415–420 (1976).
- Metzger, H. Transmembrane signaling: the joy of aggregation. *J. Immunol.* **149**, 1477–1487 (1992).
- Isakov, N. ITIMs and ITAMs. The Yin and Yang of antigen and Fc receptor-linked signaling machinery. *Immunol. Res.* **16**, 85–100 (1997).
- Gessner, J. E., Heiken, H., Tamm, A. & Schmidt, R. E. The IgG Fc receptor family. *Ann. Hematol.* **76**, 231–248 (1998).
- Galon, J. *et al.* Soluble Fcγ receptors: interaction with ligands and biological consequences. *Int. Rev. Immunol.* **16**, 87–111 (1997).
- Deisenhofer, J. Crystallographic refinement and atomic models of a human Fc fragment and its complex with fragment B of protein A from *Staphylococcus aureus* at 2.9- and 2.8 Å resolution. *Biochemistry* **20**, 2361–2370 (1981).
- Sauer-Eriksson, A. E., Kleywegt, G. J., Uhlen, M. & Jones, T. A. Crystal structure of the C2 fragment of streptococcal protein G in complex with the Fc domain of human IgG. *Structure* **3**, 265–278 (1995).
- Burmeister, W. P., Huber, A. H. & Bjorkman, P. J. Crystal structure of the complex of rat neonatal Fc receptor with Fc. *Nature* **372**, 379–383 (1994).
- Duncan, A. R., Woof, J. M., Partridge, L. J., Burton, D. R. & Winter, G. Localization of the binding site for the human high-affinity Fc receptor on IgG. *Nature* **332**, 563–564 (1988).
- Chappel, M. S. *et al.* Identification of the Fcγ receptor class I binding site in human IgG through the use of recombinant IgG1/IgG2 hybrid and point-mutated antibodies. *Proc. Natl Acad. Sci. USA*. **88**, 9036–9040 (1991).
- Lund, J. *et al.* Multiple binding sites on the CH2 domain of IgG for mouse FcγRII. *Mol. Immunol.* **29**, 53–59 (1992).
- Greenwood, J., Clark, M. & Waldmann, H. Structural motifs involved in human IgG antibody effector functions. *Eur. J. Immunol.* **23**, 1098–1104 (1993).
- Diamond, B., Bocchini, L. & Birshtein, B. K. Site of binding of IgG2b and IgG2a by mouse macrophage Fc receptors by using cyanogen bromide fragments. *J. Immunol.* **134**, 1080–1083 (1985).
- Gergely, J. & Sarmay, G. The two binding-site models of human IgG binding Fcγ receptors. *FASEB J.* **4**, 3275–3283 (1990).

15. Sondermann, P., Huber, R. & Jacob, U. Crystal structure of the soluble form of the human Fc γ -receptor IIb: a new member of the immunoglobulin superfamily at 1.7 Å resolution. *EMBO J.* **18**, 1095–1103 (1999).
16. Jefferis, R., Lund, J. & Pound, J. D. IgG-Fc-mediated effector functions: molecular definition of interaction sites for effector ligands and the role of glycosylation. *Immunol. Rev.* **163**, 59–76 (1998).
17. Woof, J. M., Partridge, L. J., Jefferis, R. & Burton, D. R. Localisation of the monocyte-binding region on human immunoglobulin G. *Mol. Immunol.* **23**, 319–330 (1986).
18. Maxwell, K. F. *et al.* Crystal structure of the human leukocyte Fc receptor, Fc γ RIIa. *Nature Struct. Biol.* **6**, 437–442 (1999).
19. Garman, S. C., Kinet, J. -P. & Jardetzky, T. S. Crystal structure of the human high-affinity IgE receptor. *Cell* **95**, 951–961 (1998).
20. Tamm, A. & Schmidt, R. E. IgG binding sites on human Fc γ receptors. *Int. Rev. Immunol.* **16**, 57–85 (1997).
21. Sarmay, G., Lund, J., Rozsnyay, Z., Gergely, J. & Jefferis, R. Mapping and comparison of the interaction sites on the Fc region of IgG responsible for triggering antibody dependent cellular cytotoxicity (ADCC) through different types of human Fc γ receptor. *Mol. Immunol.* **29**, 633–639 (1992).
22. Sondermann, P., Jacob, U., Kutscher, C. & Frey, J. Characterization and crystallization of soluble human Fc γ receptor II (CD32) isoforms produced in insect cells. *Biochemistry* **38**, 8469–8477 (1999).
23. Ghirlando, R. *et al.* Stoichiometry and thermodynamics of the interaction between the Fc fragment of human IgG1 and its low-affinity receptor Fc γ RIII. *Biochemistry* **34**, 13320–13327 (1995).
24. Edberg, J. C. & Kimberly, R. P. Cell type-specific glycoforms of Fc γ RIIIa (CD16): differential ligand binding. *J. Immunol.* **159**, 3849–3857 (1997).
25. Ishizaka, T. & Ishizaka, K. Triggering from histamine from rat mast cells by divalent antibodies against IgE receptor. *J. Immunol.* **120**, 800–806 (1978).
26. Hulett, M. D., Witort, E., Brinkworth, R. I., McKenzie, I. F. C. & Hogarth, P. M. Multiple regions of human Fc γ RII (CD32) contribute to the binding of IgG. *J. Biol. Chem.* **270**, 21188–21194 (1995).
27. Cook, J. P. *et al.* Identification of contact residues in the IgE binding site of human Fc ϵ RI α . *Biochemistry* **36**, 15579–15588 (1997).
28. Hulett, M. D., Brinkworth, R. I., McKenzie, I. F. & Hogarth, P. M. Fine structure analysis of interaction of Fc ϵ RI with IgE. *J. Biol. Chem.* **274**, 13345–13352 (1999).
29. Kato, K. *et al.* Structural basis of the interaction between IgG and Fc γ Receptors. *J. Mol. Biol.* **295**, 213–224 (2000).
30. Tax, W. J. M., Willems, H. W., Reekers, P. P. M., Capel, P. J. A. & Koene, R. A. P. Polymorphism in mitogenic effect of IgG1 monoclonal antibodies against T3 antigen on human T cells. *Nature* **304**, 445–447 (1983).
31. Warmerdam, P. A. M., van de Winkel, J. G. J., Vlug, A., Westerdaal, N. A. C., Capel, P. J. A. A single amino acid in the second Ig-like domain of the human Fc γ receptor II is critical for human IgG2 binding. *J. Immunol.* **147**, 1338–1343 (1991).
32. Hulett, M. D. & Hogarth, P. M. The second and third extracellular domains of Fc γ RI (CD64) confer the unique high affinity binding of IgG2a. *Mol. Immunol.* **35**, 989–996 (1998).
33. Sutton, B. J. & Gould, H. J. The human IgE network. *Nature* **366**, 421–428 (1993).
34. Keown, M. B., Henry, A. J., Ghirlando, R., Sutton, B. J. & Gould, H. J. Thermodynamics of the interaction of human immunoglobulin E with its high-affinity receptor Fc ϵ RI. *Biochemistry* **37**, 8863–8869 (1998).
35. Henry, A. J. *et al.* Participation of the N-terminal region of Ce3 in the binding of human IgE to its high-affinity receptor Fc ϵ RI. *Biochemistry* **36**, 15568–15578 (1997).
36. Clynes, R., Dumitru, C. & Ravetch, J. V. Uncoupling of immune complex formation and kidney damage in autoimmune glomerulonephritis. *Science* **279**, 1052–1054 (1998).
37. Clynes, R. A., Towers, T. L., Presta, L. G. & Ravetch, J. V. Inhibitory Fc receptors modulate in vivo cytotoxicity against tumor targets. *Nature Med.* **6**, 443–446 (2000).
38. Woodle, E. S. *et al.* Phase I trial of a humanized, Fc receptor nonbinding OKT3 antibody, huOKT3 γ_1 (Ala–Ala) in the treatment of acute renal allograft rejection. *Transplantation* **68**, 608–616 (1999).
39. Reddy, M. P. *et al.* Elimination of Fc receptor-dependent effector functions of a modified IgG4 monoclonal antibody to human CD4. *J. Immunol.* **164**, 1925–1933 (2000).
40. Curnow, R. T. Clinical experience with CD64-directed immunotherapy: an overview. *Cancer Immunol. Immunother.* **45**, 210–215 (1997).
41. Sondermann, P. & Jacob, U. Human Fc γ receptor IIb expressed in *Escherichia coli* reveals IgG binding capability. *Biol. Chem.* **380**, 717–721 (1999).
42. Rutishauser, U. *et al.* The covalent structure of a human gamma G-immunoglobulin. 8. amino acid sequence of heavy-chain cyanogen bromide fragments H5–H7. *Biochemistry* **9**, 3171–3181 (1970).
43. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
44. Navaza, J. AMoRe: an automated package for molecular replacement. *Acta Crystallogr. A* **50**, 157–163 (1994).
45. Brunger, A. T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
46. Kraulis, P. J. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
47. Merritt, E. A. & Murphy, M. E. P. Raster3D Version 2. 0. a program for photorealistic molecular graphics. *Acta Crystallogr. D* **50**, 869–873 (1994).
48. Barton, G. C. ALS-CRIP: a tool to format multiple sequence alignments. *Protein Eng.* **6**, 37–40 (1993).

Acknowledgements

We thank R. E. Schmidt and T. Witte and the unknown myeloma patient for providing us with the monoclonal hIgG1, G. P. Bourenkov and H. D. Bartunik, MPG-ASMB DESY/Hamburg, for help with the diffraction measurements.

Correspondence and requests for materials should be addressed to P.S. (e-mail: sonderma@biochem.mpg.de). Atomic coordinates of the Fc γ RIII (accession code: le4j) and of the hFc1–Fc γ RIII complex (accession code: le4k) have been deposited in the Protein Data Bank.

Determining the ages of comets from the fraction of crystalline dust

Joseph A. Nuth III*, Hugh G. M. Hill* & Gunther Kletetschka*†

Astrochemistry Branch, Code 691* & Code 921†, NASA's Goddard Space Flight Centre, Greenbelt, Maryland 20771, USA

The timescale for the accretion of bodies in the disk surrounding a young star depends upon a number of assumptions, but there are few observational constraints. In our own Solar System, measurements of meteoritic components can provide information about the inner regions of the nebula, but not the outer parts. Observations of the evolution of more massive protostellar systems (Herbig A_c/B_c stars) imply that significant changes occur in the physical properties of their dust with time¹. The simplest explanation is that thermal annealing of the original, amorphous grains in the hot inner nebula slowly increases the fractional abundance of crystalline material over time. Crystalline dust is then transported outward, where it is incorporated into comets that serve as a long-term reservoir for dust disks, such as that surrounding Beta Pictoris. Here we show that when applied to our own Solar System, this process can explain observed variations in both the volatile and dusty components of comets, while also providing a natural indicator of a comet's mean formation age. Studies of comets with different dust contents can therefore be used to investigate the timescales of the early Solar System.

Since the first observations of the dust disk around Beta Pictoris (a relatively old protostellar A star), it has been recognized² that these disks are unstable to radiative forces and are replenished from larger bodies such as comets. The mid-infrared spectrum of grains around Beta Pictoris³ is consistent with dust observed in Comet Halley and attributed to crystalline olivine^{4,5}. In addition, the mid-infrared spectra of Herbig A_c/B_c stars without companions evolve from those of amorphous astronomical silicate⁶ to spectra consistent with crystalline olivine as a function of increasing age¹. Observation of amorphous grains in the youngest systems eliminates the protostellar accretion shock as the source of crystalline minerals in these stars. Thermal annealing of amorphous, originally interstellar, grains in the inner regions of these nebulae provides a natural explanation for the observation of crystalline grains in dust disks and in comets if annealed dust can be transported to regions where comets might form.

Blue-shifted winds extending for more than a thousand astronomical units (AU) are often observed above and below the disks of Herbig A_c/B_c stars viewed edge-on (C. A. Grady, personal communication). Several authors^{7,8} have discussed the likelihood of outward mixing in a nebular environment. Outward mixing might occur by material exchange between adjacent turbulent convection cells within the disk⁸. Such a situation would provide a lower limit to the outward flux of nebular material. A more interesting suggestion is that larger-scale vortices resulting from non-linear momentum terms that are often neglected in nebular models would cause a higher rate of outward material flow in the disk⁸. Both mechanisms should work to some extent over the lifetime of the nebula, but one might easily expect the rate of outward transport to vary over time. Although the rate of outward mixing and the potential variation of this rate with time will determine the exact ratio of crystalline-to-total dust in the nebula, the increase in crystalline dust in the outer nebula should be monotonic.

The rate of spectral change in initially amorphous magnesium silicate smokes is extremely sensitive to annealing temperature⁹. Recent laboratory studies allow prediction of the mid-infrared spectrum of magnesium silicate dust based on its time-temperature

history¹⁰. The time required for silicate grains to anneal to crystalline olivine at 1,000 K is of the order of a few hundred days. This same transition occurs in about an hour at 1,050 K, yet requires centuries at 950 K and eons at 850 K (Table 1). Dust in comets comes in at least two varieties^{11,12}, crystalline and amorphous. Although amorphous dust could have been incorporated directly into growing comets, comets could not have formed under conditions that produced crystalline grains, nor could they survive the temperatures required to crystallize amorphous grains *in situ* during a previous perihelion passage. However, the temperatures required to crystallize the amorphous silicates observed in comets are found in the innermost regions of the solar nebula¹³.

Models of cometary chemistry^{14,15} have always had difficulty explaining the observed ratio of interstellar molecules¹⁶ (such as CO, N₂) to molecules produced in the solar nebula. Formation of these more complex materials—for example, hydrocarbons, ammonia, and so on—requires higher pressures and temperatures than those found in nebular models for regions beyond the orbits of Jupiter or Saturn. A mixture of nebular products with those produced sporadically in giant gaseous protoplanets has often been invoked to explain the observed ratios. Cycling just a small fraction of nebular gas and dust from the higher-pressure, higher-temperature inner nebula could offer a simple alternative explanation for the more complex, volatile chemistry of comets. However, operation of this transport mechanism throughout much of the history of the nebula would result in some very specific correlations between the chemistry of comets and the mid-infrared spectra of their dust, provided that comets form on timescales significantly shorter than the lifetime of the solar nebula.

The formation of comets in a minimum-mass nebula has been examined using an enhanced, one-dimensional model¹⁷ to follow the accretion of micrometre-size dust particles into kilometre-scale bodies. Comets form on a timescale of a few hundred thousand years, even neglecting several factors that might result in faster accretion rates. This model indicates that a growing comet accretes material from a large volume in the nebula, as the initial coagulation process is aided greatly by gas-drag-induced orbital migration. This migration homogenizes the material accreted into comets by giving them feeding zones from 10 to 100 AU in radius and could obscure small compositional differences in comets ending the accumulation phase at different nebular radii. Comets would form much more rapidly in a higher-mass nebula where gas-drag-induced orbital migration and accumulation due to gravitational instabilities would both be more important processes. The extent of the differences we might expect to observe in the dust content and in the chemical compositions of individual comets is dependent on the ratio of the comet accumulation timescale to the nebular lifetime. If these timescales are comparable, then all comets should look fairly similar. If the nebular lifetime were much longer than the time required to accumulate comets then we would predict substantial diversity in this population as the crystalline fraction of the dust and the complex organic content of the volatiles increases with time. These latter predictions certainly appear to be more consistent with observations of both the dust^{11,12} or gas^{18,19} content of recent comets.

Comets that formed very early in the history of the solar nebula will consist almost exclusively of amorphous silicates and unaltered

Table 1 Experimentally measured (and extrapolated) annealing times

Annealing temperature (K)	Time to crystallization (s)
1,100	1.9×10^0
1,050	4.1×10^3
1,000	1.9×10^7
950	2.2×10^{11}
900	7.1×10^{15}
850	7.7×10^{20}

interstellar ices since no processed material is yet available. The average comet formed late in nebular history will contain more hydrocarbons, ammonia and annealed dust than one formed earlier. We do not suggest that this increase will necessarily be linear; but, after accretion of fresh material from the surrounding molecular cloud ceases, the accumulation of processed gas and dust in the region of comet formation should at least be monotonic with time. The time-dependent nature of the dust and gas accreted into comets might easily obscure less significant differences in cometary chemistry—such as the potential distinction between comets accreted at ‘higher’ temperature in the Jupiter–Saturn region from those accreted in cooler zones near, or even beyond, the Uranus–Neptune region.

Older comets should be rich in CO, CO₂ and N₂ and amorphous dust while younger comets should contain an abundance of crystalline olivine, hydrocarbons, ammonia and other essential prebiotic compounds. Specifically, we predict that the fraction of crystalline dust is correlated to the ratios of hydrocarbons to CO and of ammonia/amines/amides to N₂. Figure 1 shows the vapour pressures^{20,21} of a variety of simple compounds including both interstellar grain components such as N₂, CO and CO₂, and more hydrogenated species that were synthesized in the nebula, for example, CH₄, C₂H₄, C₂H₆, C₃H₆, C₃H₈ and NH₃. It is clear that interstellar grains heated to above 50 K in the nebula will lose CO and N₂ from their mantles, whereas interstellar CO₂ could remain trapped. (The formation of clathrates in larger aggregates or the trapping of volatile gas in void spaces could preserve larger quantities of CO or N₂ than might otherwise be predicted by our simple model.) Hence CO and N₂ are both suitable indicators of interstellar volatiles in comets. Alternatively, if ammonia and most hydrocarbons synthesized in the nebula are cooled to below 150 K, they will be trapped on grain surfaces. CH₄ requires much cooler temperatures to achieve the same degree of condensation and might not condense into comets even if it were present in the nebula. Therefore CH₄ would not make a useful indicator of processed nebular gas in comets. A good proxy for the ratio of processed-to-primitive nebular gas in comets would be the ratio of C₂H₄ to CO and another would be the ratio of C₂H₆ to CO. The hydrocarbons represent materials synthesized primarily in the nebula and then trapped on icy grains, whereas CO is associated with grain mantles formed within a giant molecular cloud core that has never been sufficiently heated to vaporize. Using similar reasoning, the ratio of NH₃ to N₂ should also be a good measure of the ratio of processed-

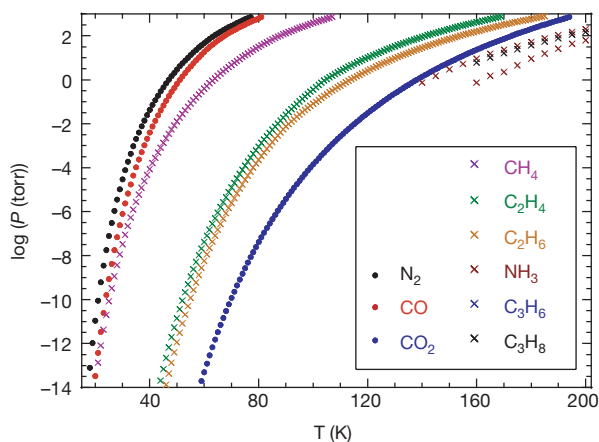


Figure 1 Vapour pressures in the range 25–200 K for a selection of simple compounds based on refs 20 and 21. The compounds include both interstellar grain mantle components (N₂, CO, CO₂) and more hydrogenated species that were almost certainly formed in the nebula (CH₄, C₂H₄, C₂H₆, C₃H₆, C₃H₈, NH₃) at relatively high pressures and temperatures.

to-primitive gas in comets. Hence, we predict that the ratio of crystalline-to-total cometary dust will be positively correlated to the ratios of C₂H₄/CO, C₂H₆/CO and NH₃/N₂ in cometary comae.

An ability to sequence comets according to their relative ages would be a useful tool in understanding a number of events in the primitive solar nebula. For example, the question of whether Jupiter formed early in the history of the nebula might be answered by comparison of the ‘age’ distribution of Oort-cloud comets versus the relative ages of comets in the Kuiper belt. Because Kuiper-belt comets formed in place they should show a wide variety of ages, whereas Oort-cloud comets might not include any of the most primitive (oldest) comets if the giant planets formed very late in nebular history. A measure of the relative accumulation age of individual comets would also provide an interesting discriminator for use in choosing targets for space flight missions. Sample returns from comets of various ages could be dated to yield the lifetime of the solar nebula. Missions to ‘old’ comets would return interstellar ices and primitive grains, whereas missions to younger comets would yield information on chemical processes in the nebula and the formation of biogenic molecules essential for the chemical evolution of life in the Solar System. Finally, the observed diversity in the chemical composition of comets could help to constrain the timescale for the formation of individual comets, compared to the lifetime of the nebula. This, in turn, could be used to provide new constraints on the mass of the solar nebula necessary to form comets at the requisite rate. □

Received 7 January; accepted 23 May 2000.

- Grady, C. A. *et al.* *Protostars and Planets IV* (eds Mannings, V., Boss, A. P. & Russell, S. S.) 613–638 (Univ. Arizona Press, Tucson, 1999).
- Smith, B. A. & Terrile, R. J. A circumstellar disk around β Pictoris. *Science* **226**, 1421–1424 (1984).
- Knacke, R. F. *et al.* The silicates in the disk of β Pictoris. *Astrophys. J.*, **418**, 440–450 (1993).
- Campins, H. & Ryan, E. V. The identification of crystalline olivine in cometary silicates. *Astrophys. J.* **341**, 1059–1066 (1989).
- Ryan, E. V. & Campins, H. Comet Halley: Spatial and temporal variability of the silicate emission feature. *Astron. J.* **101**, 695–701 (1991).
- Draine, B. T. & Lee, H. M. Optical properties of interstellar graphite and silicate grains. *Astrophys. J.* **285**, 89–108 (1984).
- Prinn, R. G. On neglect of non-linear momentum terms in solar nebula accretion disk models. *Astrophys. J.* **348**, 725–729 (1990).
- Stevenson, D. J. Chemical heterogeneity and imperfect mixing in the solar nebula. *Astrophys. J.* **348**, 730–737 (1990).
- Hallenbeck, S. L., Nuth, J. A. & Daukantas, P. L. Mid-infrared spectral evolution of amorphous magnesium silicate smokes annealed in vacuum: Comparison to cometary spectra. *Icarus* **118**, 223–270 (1995).
- Hallenbeck, S. L., Nuth, J. A. & Nelson, R. N. Evolving optical properties of annealing silicate grains: From amorphous condensate to crystalline mineral. *Astrophys. J.* **535**, 247–255 (2000).
- Hanner, M. H. The silicate material in comets. *Space Sci. Rev.* **90**, 99–108 (2000).
- Hanner, M. H., Lynch, D. K. & Russell, R. W. The 8–13 micron spectra of comets and the composition of silicate grains. *Astrophys. J.* **425**, 274–285 (1994).
- Woolum, D. S. & Cassen, P. Astronomical constraints on nebular temperatures: Implications for planetesimal formation. *Meteor. Planet. Sci.* **34**, 897–907 (1999).
- Prinn, R. G. & Fegley, B. in *Origin and Evolution of Planetary and Satellite Atmospheres* (eds Atreya, S. K., Pollock, J. B. & Matthews, M. S.) 78–136 (Univ. Arizona Press, Tucson, 1989).
- Fegley, B. in *The Chemistry of Life's Origins* (eds Greenberg, J. M., Mendoza-Gomez, C. X. & Pirronello, V.) 75–147 (Kluwer, Dordrecht, 1993).
- Gerakines, P. A. *et al.* Observations of solid carbon dioxide in molecular clouds with the Infrared Space Observatory. *Astrophys. J.* **522**, 357–377 (1999).
- Weidenschilling, S. J. The origin of comets in the solar nebula: a unified model. *Icarus* **127**, 290–306 (1997).
- A'Hearn, M. F., Millis, R. L., Schleicher, D. G., Osip, D. J. & Birth, P. V. The ensemble properties of comets: results from narrowband photometry of 85 comets, 1976–1992. *Icarus* **118**, 223–270 (1995).
- Mumma, M., Weissman, P. & Stern, S. A. in *Protostars and Planets III* (eds Levy, E. & Lunine, J.) 1177–1252 (Univ. Arizona Press, Tucson, 1993).
- Brown, G. N. & Ziegler, W. T. in *Advances in Cryogenic Engineering* Vol. 25 (eds Timmerhaus, K. D. & Snyder, H. A.) 662–670 (Plenum, New York, 1980).
- Lide, David R. Lide (ed.) *CRC Handbook of Chemistry and Physics: 75th edn* Section 6, 66–76 (CRC Press, Boca Raton, 1994).

Acknowledgements

H.G.M.H. & G.K. acknowledge support for this work received under the NAS/NRC Resident Research Associate Program at GSFC. The authors also thank F. Ferguson, D. Fixsen and J. Allen for discussion and technical help with the preparation of the manuscript.

Correspondence and requests for materials should be addressed to J.A.N. (e-mail: uljan@lepvax.gsfc.nasa.gov).

Gain-assisted superluminal light propagation

L. J. Wang, A. Kuzmich & A. Dogariu

NEC Research Institute, 4 Independence Way, Princeton, New Jersey 08540, USA

Einstein's theory of special relativity and the principle of causality^{1–4} imply that the speed of any moving object cannot exceed that of light in a vacuum (c). Nevertheless, there exist various proposals^{5–18} for observing faster-than- c propagation of light pulses, using anomalous dispersion near an absorption line^{4,6–8}, nonlinear⁹ and linear gain lines^{10–18}, or tunnelling barriers¹⁹. However, in all previous experimental demonstrations, the light pulses experienced either very large absorption⁷ or severe reshaping^{9,19}, resulting in controversies over the interpretation. Here we use gain-assisted linear anomalous dispersion to demonstrate superluminal light propagation in atomic caesium gas. The group velocity of a laser pulse in this region exceeds c and can even become negative^{16,17}, while the shape of the pulse is preserved. We measure a group-velocity index of $n_g = -310(\pm 5)$; in practice, this means that a light pulse propagating through the atomic vapour cell appears at the exit side so much earlier than if it had propagated the same distance in a vacuum that the peak of the pulse appears to leave the cell before entering it. The observed superluminal light pulse propagation is not at odds with causality, being a direct consequence of classical interference between its different frequency components in an anomalous dispersion region.

When a light pulse of frequency ν and bandwidth $\Delta\nu$ enters a dispersive linear medium of an optical refractive index $n(\nu)$, the light pulse propagates at the group velocity $v_g = c/n_g$, where

$n_g = n(\nu) + \nu dn(\nu)/d\nu$ is the group-velocity index. If the group-velocity index remains constant over the pulse bandwidth $\Delta\nu$, the light pulse maintains its shape during propagation. In recent experiments involving electromagnetically induced transparency (EIT)^{20–22}, the group-velocity index was greatly enhanced using the lossless normal dispersion region between two closely spaced absorption lines. Thus the group velocity of light was dramatically reduced to as slow as 8 m s^{-1} (refs 23–25). Conversely, between two closely spaced gain lines¹⁶, an anomalous dispersion region appears where $\nu dn(\nu)/d\nu$ is negative and its magnitude can become large. In this situation, the group velocity of a light pulse can exceed c and can even become negative^{16,17}.

A negative group velocity of light is counterintuitive but can be understood as follows. For a medium of a length L , it takes a propagation time $L/v_g = n_g L/c$ for a light pulse to traverse it. Compared with the propagation time for light to traverse the same distance in a vacuum, that is, the vacuum transit time L/c , the light pulse that enters the medium will exit at a moment that is delayed by a time difference $\Delta T = L/v_g - L/c = (n_g - 1)L/c$. When $n_g < 1$, the delay time ΔT is negative, resulting in an advancement. In other words, when incident on a medium with group-velocity index $n_g < 1$, a light pulse can appear on the other side sooner than if it had traversed the same distance in a vacuum^{5,15}. Furthermore, in contradiction to traditional views that a negative group velocity of light has no physical meaning, when the group-velocity index becomes negative, the pulse advancement $-\Delta T = (1 - n_g)L/c$ becomes larger than the vacuum transit time L/c . In other words, it appears as if the pulse is leaving the cell even before it enters. This counterintuitive phenomenon is a consequence of the wave nature of light.

The principle of the experimental realization of a lossless anomalous dispersion region and hence gain-assisted superluminality (GAS) is illustrated in Fig. 1a. In a gaseous medium of atoms each of which has three levels: an excited state $|0\rangle$ and two ground states $|1\rangle$ and $|2\rangle$, we first prepare all atoms to be in a ground state $|1\rangle$ by optical pumping. For simplicity, we first ignore the Doppler shift and assume that the atoms are at rest. We apply two strong continuous-wave (CW) Raman pump light beams E_1 and E_2 that propagate through the atomic medium. The frequencies of E_1 and E_2 , ν_1 and ν_2 , are different by a small amount 2Δ and both fields are detuned from the atomic transition frequency ν_{01} ($|0\rangle$ to $|1\rangle$) by a large average amount Δ_0 . Since the Rabi frequencies associated with the fields E_1 and E_2 are small compared with the common detuning Δ_0 , the atoms mostly remain in state $|1\rangle$. When a probe light beam E_p is introduced, a Raman transition can occur, causing an atom to absorb a Raman pump photon from the fields E_1 or E_2 and emit a photon into the field E_p while making a transition from $|1\rangle$ to $|2\rangle$. Obviously, there are two frequencies where the gain in the probe field is maximized. The maximum gain occurs when the probe field is resonant with the Raman transitions caused by either of the two pump fields E_1 and E_2 . The optical susceptibility of the probe field can thus be derived as

$$\chi(\nu) = \frac{M_1}{\nu - \nu_1 + i\gamma} + \frac{M_2}{\nu - \nu_2 + i\gamma} \quad (1)$$

Here $M_{1,2} = \frac{\mu_{02}^2 |\Omega_{1,2}|^2}{4\pi\hbar\epsilon_0 \Delta_0^2} N$ with μ_{02} , $\Omega_{1,2}$ and N being the dipole moment of the $|0\rangle$ to $|2\rangle$ atomic transition, the Rabi frequencies of the Raman pump fields E_1 and E_2 and the effective atomic density difference of states $|1\rangle$ and $|2\rangle$, respectively. γ is the Raman transition inverse lifetime. The refractive index and the gain coefficient obtained using the susceptibility given in equation (1) are shown in Fig. 1b. In the region between the two gain lines, an anomalous dispersion region appears.

However, in a gaseous atomic medium, there is Doppler broadening: for atoms moving at different velocities in the light propagation direction, the common detuning Δ_0 is shifted. The effects of Doppler broadening are twofold. First, the M -factors in equation

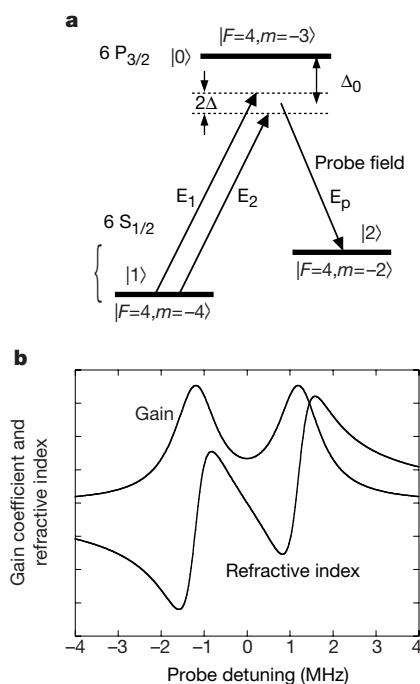


Figure 1 Gain-assisted anomalous dispersion. **a**, Schematic atomic level diagram. **b**, Frequency-dependent gain coefficient and refractive index obtained from equation (1) for a probe light beam propagating through an atomic medium with its level structure shown in **a**.

(1) are replaced by

$$M_{1,2} = \frac{|\mu_{02}|^2}{4\pi\hbar\epsilon_0} \int dV \frac{|\Omega_{1,2}|^2}{(\Delta_0 + \nu V/c)^2} N(V) \quad (2)$$

Here $N(V)$ is the effective density of atoms in the velocity group V . The common detuning Δ_0 is replaced by the Doppler detuning $\Delta_0 + \nu V/c$. However, the quadratic dependence of the coefficient M on the Doppler detuning prevents cancellation. Second, inside the Doppler profile, the expression for M given above appears to become singular when the detuning $\Delta_0 + \nu V/c$ vanishes. This is automatically avoided in practice because for atoms with detuning $\Delta_0 + \nu V/c$ that is less than a certain linewidth containing contributions from the natural linewidth and power broadening, the Raman pump beams act like reversed optical pumping beams that deplete these velocity groups. For atoms in the velocity group where the effective detuning $\Delta_0 + \nu V/c$ vanishes, $N(V)$, the atom number difference also vanishes. Finally, the atoms that are pumped away from the level $|1\rangle$ act as a weak broadband absorber. This compensates for the small residual gain in the region between the two Raman gain lines shown in Fig. 1b.

The experiment was performed using an atomic caesium (Cs) vapour cell at 30 °C and the main set-up is shown in Fig. 2. The caesium atoms are confined in a 6-cm-long Pyrex glass cell coated with paraffin for the purpose of maintaining ground-state spin polarization. The atomic cell is placed in a small (1.0 G) uniform magnetic field parallel to the light propagation direction. In region I, two laser beams optically pump the atoms into the ground-state hyperfine magnetic sub-level $|F = 4, m = -4\rangle$ that serves as state $|1\rangle$ (Fig. 1a). One left-hand (σ^-) polarized laser beam from a narrow linewidth diode laser is tuned to the 852-nm D_2 transitions to empty the $6S_{1/2} F = 3$ hyperfine ground state. We further apply a second laser (σ^-) to optically pump the atoms into the $|F = 4, m = -4\rangle$ state via the D_1 transitions to the $6P_{1/2}$ hyperfine excited states. When atoms collide with the paraffin-coated glass walls, they change their velocities inside the Doppler profile while remaining in the ground state $|F = 4, m = -4\rangle$ and hence the majority of the caesium atoms inside the cell are prepared into this state. In region II, three light beams derived from the same laser propagate co-linearly through the cell. Two strong CW Raman pump beams are right-hand circularly polarized (σ^+) and are frequency-shifted by

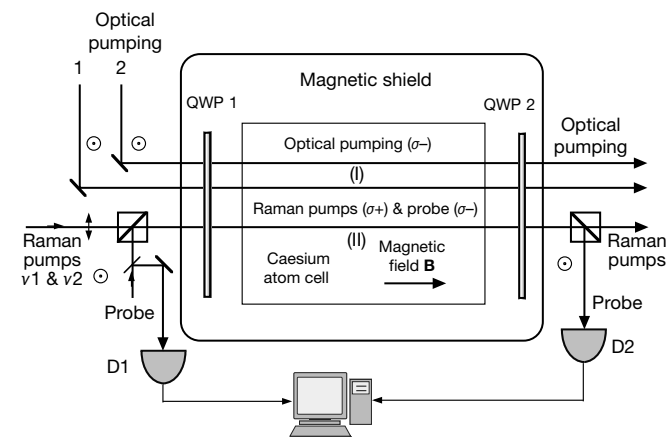


Figure 2 Schematic experimental set-up. Two optical pumping beams tuned to the caesium (Cs) atomic D_1 and D_2 transitions prepare the atoms in its ground-state hyperfine sublevel $|F = 4, m = -4\rangle$. Two Raman pump beams and a Raman probe beam derived from a common narrow linewidth diode laser propagate co-linearly parallel to a small magnetic field $\mathbf{B}$ through the atomic cell. Two $\lambda/4$ plates (QWP1&2) are used to prepare the three light beams into the corresponding circular polarization states and then separate them for analysis.

2.7 MHz using two acousto-optical (A/O) modulators. The linewidths of the A/O modulators are 20 kHz. A third light beam, the probe beam, is left-hand polarized (σ^-) and by using another A/O can be tuned in frequency and operate either in CW or pulsed mode. In the experiment, the pre-empted hyperfine magnetic sublevel $|F = 4, m = -2\rangle$ serves as the Raman transition final state $|2\rangle$. The intermediate Raman transitional state $|0\rangle$ is served primarily by the hyperfine sublevel $|F = 4, m = -3\rangle$ of the $6P_{3/2}$ excited state with additional contributions from transitions through $|F = 3, m = -3\rangle$ and $|F = 5, m = -3\rangle$ hyperfine sublevels.

First, we operate the Raman probe beam in a tunable CW mode to measure the gain and refractive index of the atomic system as a function of the probe frequency detuning. Figure 3 shows the measured gain coefficient and the refractive index. In order to obtain the gain coefficient, we first measure the intensity of the transmitted probe beam as a function of probe frequency. We then extract the gain coefficient. The refractive index is measured using a radio-frequency interferometric technique. The superimposed curve is obtained from equation (1) using parameters obtained from the gain measurement. From Fig. 3, we see that a negative change of $\Delta n = -1.8 \times 10^{-6}$ in the index occurs over a narrow probe frequency range of $\Delta\nu = 1.9$ MHz between the two gain lines. Using the expression of the group-velocity index, we obtain the result $n_g = -330 (\pm 30)$ in that frequency region. The 10% error reflects the accuracy of the phase measurement.

Next, a pulsed Raman probe beam is employed to observe the superluminal propagation. A near-gaussian probe pulse with a 3.7- μ s full-width at half-maximum (FWHM) is generated by applying an electronic pulse to the probe beam A/O modulator. A portion of the pulsed probe beam is divided at a beam-splitter before the atomic cell and aligned onto photodiode D1 as a reference. Because the total number of atoms in the probe volume limits the maximum net energy gain of the probe pulse, we use a very weak probe beam ($<1 \mu$ W) in order to avoid saturation and hence to optimize the anomalous dispersion. A high-sensitivity avalanche photodiode, reverse-biased below breakdown, serves as detector D2 to measure the weak probe pulse that propagates through the atomic cell. The photoelectric current produced by detector D2 is converted to a voltage signal using a 500- Ω load resistor and recorded by a digitizing oscilloscope using a synchronized output signal from the pulse generator as the trigger. Pulses from detector D1 are also recorded.

In order to measure the pulse propagation time, we first tune the diode laser that produces the Raman pump and probe beams far off-resonance from the 852-nm caesium D_2 lines (by 2.5 GHz) to

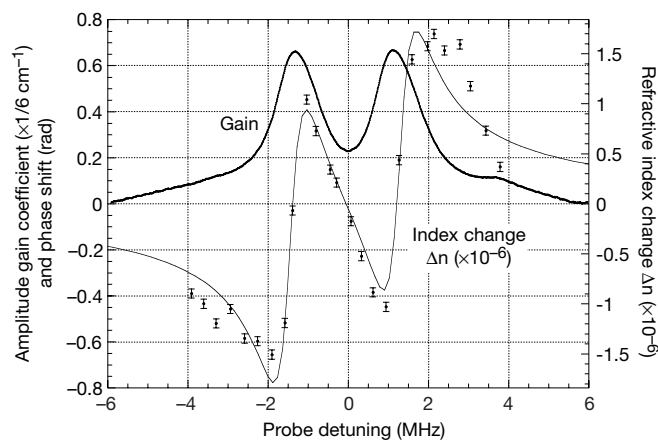


Figure 3 Measured refractive index and gain coefficient. The superimposed curve over the index data is obtained using equation (1) with parameters ν_1 , ν_2 and γ obtained experimentally.

measure the time-dependent probe-pulse intensity. When the laser is placed far off-resonance, the atoms have no effect and the probe pulse propagates at speed c inside the cell. We then tune the diode laser back to within the Doppler absorption profile and lock the laser on its side. Using the same synchronized pulse generator output signal as the trigger, we record the time-dependent probe-pulse intensity measured by detector D2. We verify that no systematic drift is present by tuning the laser off-resonance again by the same amount and record the probe-pulse signal; the two off-resonance pulses are identical to within less than 1 ns. Probe pulses both on and off-resonance are shown in Fig. 4. Probe pulses on resonance show a 40% transmittance and this is due to the broadband absorption of those atoms reverse-pumped away from the $|F = 4, m = -4\rangle$ state. It is evident that there is almost no change in the pulse shape. The front edges and the trailing edges of the pulses are shown in the insets; both edges are shifted forward by the same amount. Using a least-squares fitting procedure, we obtain a pulse advancement shift of $62 (\pm 1)$ ns. Compared with the 0.2-ns propagation time for light to traverse the 6-cm length of the atomic cell in vacuum, the 62-ns advancement gives an effective group-velocity index of $n_g = -310 (\pm 5)$. The small discrepancy with the group-velocity index inferred from the refractive index data is due to experimental errors. The pulses measured with detector D1 are also recorded in the sequence of the off-, on-, off-resonance pulse-propagation measurements and are found to be identical to within 1.5 ns.

Here we note that the physical mechanism that governs the observed superluminal light propagation differs for the previously studied anomalous dispersion associated with an absorption or a gain resonance^{5–17}. Specifically, in the anomalous dispersion region of a single gain resonance, the superluminal propagation of a pulse has been viewed as the result of the amplification of the pulse front edge and absorption of its tail¹. In the present experiment, the 3.7- μ s FWHM probe pulse has only a 120-kHz bandwidth (FWHM) that is much narrower than the 2.7-MHz separation of the two gain lines spectrally. The probe pulse is placed in the middle of these gain lines. The probe pulse thus contains essentially no spectral components that are resonant with the Raman gain lines to be amplified. Therefore, the argument that the probe pulse is advanced by amplification of its front edge does not apply. The superluminal light propagation observed here is the result only of the anomalous

dispersion region created with the assistance of two nearby Raman gain resonances. We emphasize that the observed superluminal light propagation is a result of the wave nature of light². It can be understood by the classical theory of wave propagation in an anomalous dispersion region where interference between different frequency components produces this rather counterintuitive effect.

Finally, we note that the observed superluminal light pulse propagation is not at odds with causality or special relativity. In fact, the very existence of the lossless anomalous dispersion region given in equation (1) is a result of the Kramers–Kronig relation which itself is based on the causality requirements of electromagnetic responses^{3,5}. Remarkably, the signal velocity⁴ of a light pulse, defined as the velocity at which the half point of the pulse front travels, also exceeds the speed of light in a vacuum, c , in the present experiment. It has also been suggested^{4,16} that the true speed at which information is carried by a light pulse should be defined as the “frontal” velocity of a step-function-shaped signal which has been shown not to exceed c (ref. 4). The implications of the present experiment on signal propagation and its speed will be further analysed, particularly for the case when the light pulse consists of only a few photons. □

Received 11 May; accepted 26 June 2000.

1. Einstein, A., Lorentz, H. A., Minkowski, H. & Weyl, H. *The Principle of Relativity, Collected Papers* (Dover, New York, 1952).
2. Born, M. & Wolf, E. *Principle of Optics* 7th edn (Cambridge Univ. Press, Cambridge, 1997).
3. Landau, L. D. & Lifshitz, E. M. *Electrodynamics of Continuous Media* (Pergamon, Oxford, 1960).
4. Brillouin, L. *Wave Propagation and Group Velocity* (Academic, New York, 1960).
5. Chiao, R. Y. in *Amazing Light, a Volume Dedicated to C. H. Townes on His Eightieth Birthday* (ed. Chiao, R. Y.) 91–108 (Springer, New York, 1996).
6. Garrett, C. G. B. & McCumber, D. E. Propagation of a gaussian light pulse through an anomalous dispersion medium. *Phys. Rev. A* **1**, 305–313 (1970).
7. Chu, S. & Wong, S. Linear pulse propagation in an absorbing medium. *Phys. Rev. Lett.* **48**, 738–741 (1982).
8. Akulshin, A. M., Barreiro, S. & Lezama, A. Steep anomalous dispersion in coherently prepared Rb vapor. *Phys. Rev. Lett.* **83**, 4277–4280 (1999).
9. Basov, N. G., Ambartsumyan, R. V., Zuev, V. S., Kryukov, P. G. & Letokhov, V. S. Nonlinear amplification of light pulses. *Sov. Phys. JETP* **23**, 16–22 (1966).
10. Casperson, L. & Yariv, A. Pulse propagation in a high-gain medium. *Phys. Rev. Lett.* **26**, 293–295 (1971).
11. Içsevgi, A. & Lamb, W. E. Propagation of light pulses in a laser amplifier. *Phys. Rev.* **185**, 517–545 (1969).
12. Picholle, E., Montes, C., Leycuras, C., Legrand, O. & Botineau, J. Observation of dissipative superluminal solitons in a Brillouin fiber ring laser. *Phys. Rev. Lett.* **66**, 1454–1457 (1991).
13. Fisher, D. L. & Tajima, T. Superluminal laser pulse in an active medium. *Phys. Rev. Lett.* **71**, 4338–4341 (1993).
14. Chiao, R. Y. Superluminal (but causal) propagation of wave packets in transparent media with inverted atomic populations. *Phys. Rev. A* **48**, R34–R37 (1993).
15. Bolda, E. L., Chiao, R. Y. & Garrison, J. C. Two theorems for the group velocity in dispersive media. *Phys. Rev. A* **48**, 3890–3894 (1993).
16. Steinberg, A. M. & Chiao, R. Y. Dispersionless, highly superluminal propagation in a medium with a gain doublet. *Phys. Rev. A* **49**, 2071–2075 (1994).
17. Mitchell, M. W. & Chiao, R. Y. Causality and negative group delays in a simple bandpass amplifier. *Am. J. Phys.* **66**, 14–19 (1998).
18. Bolda, E., Garrison, J. C. & Chiao, R. Y. Optical pulse propagation at negative group velocities due to a nearby gain line. *Phys. Rev. A* **49**, 2938–2947 (1994).
19. Steinberg, A. M., Kwiat, P. G. & Chiao, R. Y. Measurement of the single-photon tunneling time. *Phys. Rev. Lett.* **71**, 708–711 (1993).
20. Harris, S. E. Electromagnetically induced transparency. *Phys. Today* **50**, 36–42 (1997).
21. Scully, M. O. & Zubairy, M. S. *Quantum Optics* (Cambridge Univ. Press, 1997).
22. Xiao, M., Li, Y.-Q., Jin, S.-Z. & Gea-Banacloche, J. Measurement of dispersion properties of electromagnetically induced transparency in rubidium atoms. *Phys. Rev. Lett.* **74**, 666–669 (1995).
23. Hau, L. V., Harris, S. E., Dutton, Z. & Behroozi, C. H. Light speed reduction to 17 meters per second in an ultracold atomic gas. *Nature* **397**, 594–598 (1999).
24. Kash, M. M. *et al.* Ultraslow group velocity and enhanced nonlinear optical effects in a coherently driven hot atomic gas. *Phys. Rev. Lett.* **82**, 5229–5232 (1999).
25. Budker, D., Kimball, D. F., Rochester, S. M. & Yashchuk, V. V. Nonlinear magneto-optics and reduced group velocity of light in atomic vapor with slow ground state relaxation. *Phys. Rev. Lett.* **83**, 1767–1770 (1999).

Acknowledgements

We thank R. A. Linke for several stimulating discussions. We thank D. K. Walter, W. Happer, J. A. Giordmaine, D. J. Chadi, S. A. Solin, R. Y. Chiao, S. E. Harris and E. S. Polzik for helpful discussions. We thank E. B. Alexandrov and N. P. Bigelow for the use of the paraffin-coated caesium cells.

Correspondence and requests for materials should be addressed to L.J.W. (e-mail: Lwan@research.nj.nec.com).

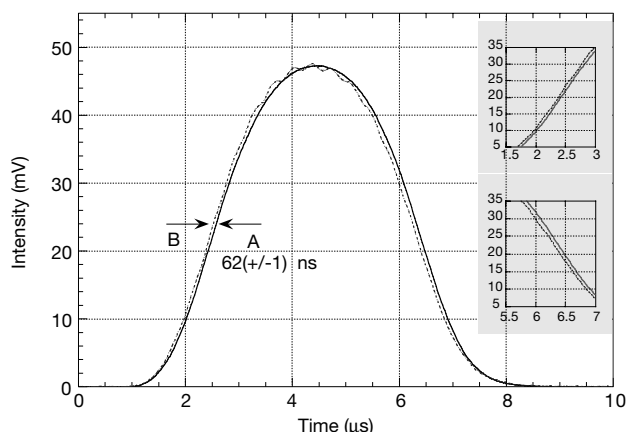


Figure 4 Measured pulse advancement for a light pulse traversing through the caesium vapour in the gain-assisted superluminality state. A indicates a light pulse far off-resonance from the caesium D₂ transitions propagating at speed c through 6 cm of vacuum. B shows the same light pulse propagating through the same caesium-cell near resonance with a negative group velocity $-c/310$. Insets show the front and trailing portions of the pulses. Pulses A and B are both the average of 1,000 pulses. The off-resonance pulse (A) is normalized to the magnitude of B.

Non-collinear states in magnetic sensors

Adrian Taga, Lars Nordström, Peter James, Börje Johansson & Olle Eriksson

Condensed Matter Theory Group, Department of Physics, University of Uppsala, Box 530, S-751 21, Uppsala, Sweden

Certain materials have an electrical conductivity that is extremely sensitive to an applied magnetic field; this phenomenon, termed 'giant magnetoresistance'^{1–3}, can be used in sensor applications. Typically, such a device comprises several ferromagnetic layers, separated by non-magnetic spacer layer(s)—a so-called 'super-lattice' geometry^{1–3}. In the absence of a magnetic field, the ferromagnetic layers may be magnetized in opposite directions by interlayer exchange coupling, while an applied external magnetic field causes the magnetization directions to become parallel. Because the resistivity depends on the magnetization direction, an applied field that changes the magnetic configuration may be detected simply by measuring the change in resistance. In order to detect weak fields, the energy difference between different magnetization directions should be small; this is usually achieved by using many non-magnetic atomic spacer layers. Here we show, using first-principles theory, that materials combinations such as Fe/V/Co multilayers can produce a non-collinear magnetic state in which the magnetization direction between Fe and Co layers differs by about 90°. This state is energetically almost degenerate with the collinear magnetic states, even though the number of non-magnetic vanadium spacer layers is quite small.

When electrons travel through a ferromagnetic super-lattice they experience an effective potential that depends on the spin direction. The spin direction in the magnetic layers is different from that in the non-magnetic layers. The spin-down electrons have a higher energy barrier to pass through than the spin-up electrons and this gives

rise to a lowered transport coefficient for the spin-down states. However, owing to the good conductance of the spin-up electrons, the total electronic transport is high. In the case of anti-parallel coupling, spin-up and spin-down electrons experience a high effective potential in alternating magnetic layers, which reduces the electrical transport properties of both spin channels so that the total electrical resistance is increased dramatically. This difference in resistance can be used to detect small fields for use in, for instance, sensor applications^{1–3}.

To detect very weak magnetic fields, the interaction between the magnetic layers must be weak. This is achieved by increasing the width of the non-magnetic spacer layer. This layer typically has a thickness of 20–40 atomic layers^{1–3}. However, an increased width of this non-magnetic spacer layer will decrease the number of magnetic layers per unit volume. As it is the presence of magnetic layers which produces the difference in resistance between the ferro- and anti-ferromagnetic configurations, it is desirable to increase the number of magnetic layers within a certain volume.

In addition, non-epitaxial growth and increased lattice imperfections (such as stacking faults) occur in traditional devices; these diminish the giant magnetoresistance (GMR) effect. This effect is caused by the fact that the non-magnetic and magnetic layers have different lattice constants and the strain field produced by the lattice mismatch causes imperfections in the crystal structure. The lattice imperfections give rise to incoherent scattering, which may destroy the spin information of the electrons, which in turn reduces and degrades the GMR effect. The thicker the intermediate non-magnetic layers, the more lattice imperfections occur. Therefore, it is also for this reason important to make the non-magnetic layers as thin as possible, while still keeping the interlayer exchange coupling weak, to enable detection of tiny magnetic fields.

Here we demonstrate that, by suitable tuning of the super-lattice geometry and an appropriate choice of materials, we can stabilize a new magnetic configuration. The successive magnetic layers do not couple in a simple ferro- or anti-ferromagnetic way, but in an essentially perpendicular, non-collinear fashion; every second magnetic layer has a magnetization direction pointing 'out-of-plane', as opposed to the 'in-plane' intermediate layers. This unusual

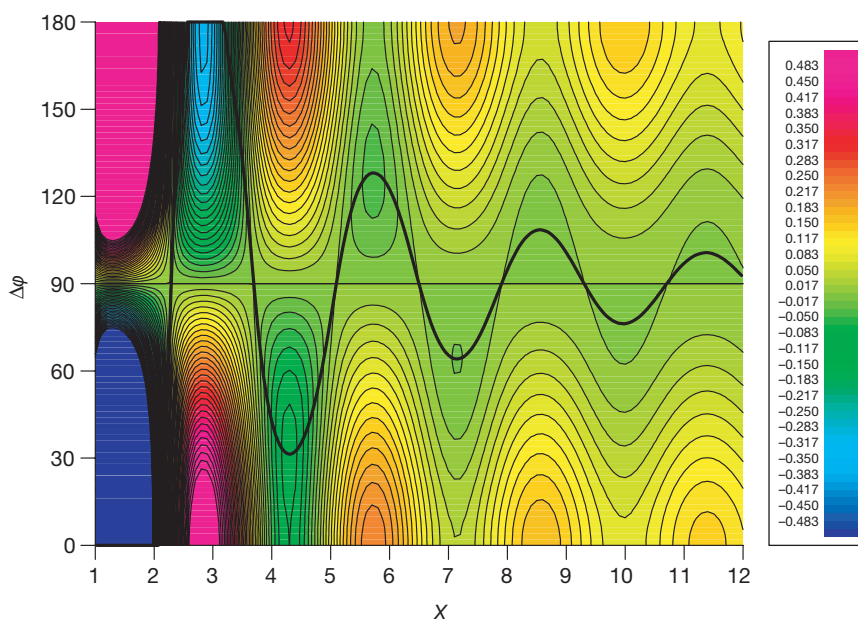


Figure 1 Calculated values of the difference in magnetization directions of Fe and Co as a function of number of V spacer layers (x). The energy landscape is shown as contours and in colour code as indicated on the right-hand side of the figure. The calculations made here employed the linear muffin-tin orbital (LMTO) method and the Kohn–Sham equations

are solved for a general potential without any shape approximation. The calculations were well converged with respect to all parameters involved, such as k -space sampling, basis set truncation and total energy. Details of the method can be obtained from J. M. Wills (unpublished work).

magnetic configuration is shown to have improved properties for use in sensor applications, even when the number of non-magnetic spacer layers is small. Hence the above-mentioned problem of incoherent scattering and lattice imperfections, caused by the strain field, is greatly reduced.

The geometry of the magnetic super-lattice discussed here consists of alternating iron (Fe), vanadium (V) and cobalt (Co) layers grown in the (001) crystallographic direction, but, as will be discussed below, we can also use other materials combinations and growth directions. The sequence of atomic layers considered here is repeated to give a multilayer with the following sequence: ...Fe/V/Co/V/Fe/V/Co... The first ferromagnetic layer is Fe, epitaxially grown on V. The Fe layer (consisting of several atomic planes of Fe) has a tetragonally distorted crystal structure, caused by the lattice mismatch with the V layer (that also consists of several atomic planes), and exhibits a magnetization parallel to the layer (in-plane). The in-plane magnetization of the Fe layer is an effect of the relativistic spin-orbit coupling (resulting in magneto-crystalline anisotropy, MAE) in combination with the structural strain. (The strain was calculated to give a c/a ratio of 0.84 for Fe on V, assuming that all the atomic layers of the multi-layer adopt the in-plane lattice parameter of V and that the volume of Fe and Co is conserved and is the same as in the bulk.) Our calculations from first principles of the energy difference between in-plane and out-of-plane magnetization for strained Fe on V is 0.6 meV per Fe atom.

For the second ferromagnetic layer a Co layer (with several atomic planes) epitaxially grown on the V layer is used. The Co layer also has a tetragonally distorted crystal structure due to the lattice mismatch with the V spacer layer (calculated to give a c/a ratio of 0.80). For Co, the magneto-elastic coupling produces a magnetization parallel to the growth direction (out-of-plane) and perpendicular to the magnetization direction of the Fe layer. The calculations from first principles of the energy difference between in-plane and out-of-plane magnetizations for strained Co is 1.4 meV per Co atom; that is, more than twice as large as the value for Fe.

If there were no other physical effects, the magnetization would alternate between the Fe and the Co layers, pointing in-plane for the Fe layer and out-of-plane for the Co layer; this is a configuration we refer to as a perpendicular magnetization, in contrast to the simple ferro- or anti-ferromagnetic configuration. However, the ordinary bilinear interlayer exchange interaction between the two ferromagnetic layers (Fe and Co) tends to turn the magnetization and arrange them in parallel, in a ferromagnetic or anti-ferromagnetic arrangement⁴. If the thickness of the spacer layer is tuned, the interlayer exchange coupling can be made to be of the same size as the magneto-crystalline contribution. The two competing interactions can thus be made to produce complex non-collinear

magnetization profiles, with a tiny energy difference between different magnetic orientations.

To quantify our analysis we have calculated the strength of the interlayer exchange coupling of the proposed Fe/V/Co multilayer (with 3 Fe and 3 Co atoms). This was done in the same way as for the calculations of the MAE, that is, from first principles. The interlayer exchange coupling was found to oscillate in a RKKY-like fashion as a function of V thickness; the calculated exchange coupling was fitted to a functional form $E_{\text{ex}} = I(x)\cos(\phi_{\text{Fe}} - \phi_{\text{Co}})$ where $I(x)$ is the strength of the interlayer exchange coupling, x is the number of V layers, and $\phi_{\text{Fe(Co)}}$ is the angle between the magnetization axis for the Fe(Co) layers and the normal of the Fe(Co) layer. Calculations of this kind almost always give the correct oscillatory behaviour with correct period, but the magnitude of $I(x)$ is over-estimated by a factor of 5–10 (ref. 4). Hence, we have scaled our calculated interlayer exchange coupling by a factor of 5. (We will discuss the consequences of different scalings below; we argue that they do not influence the basic physics involved here.) We can now combine the different energy terms to obtain a simple expression for the total energy of the multilayer as a function of the angles of magnetization of the Fe and Co layer. The MAE energies are parameterized in a traditional way, as $E_{\text{MAE-Fe}} = A\cos^2(\phi_{\text{Fe}})$ for Fe and $E_{\text{MAE-Co}} = B\cos^2(\phi_{\text{Co}})$ for Co. The total energy can then be written as $E_{\text{total}} = E_{\text{ex}} + E_{\text{MAE-Fe}} + E_{\text{MAE-Co}}$. By minimizing the energy as a function of angles ϕ_{Fe} and ϕ_{Co} we calculate the magnetization directions of the Fe and Co layers as a function of the thickness of the V spacer layer.

As it is the difference in magnetization directions that is of interest, we show in Fig. 1 the calculated difference in angle, $\Delta\phi = \phi_{\text{Fe}} - \phi_{\text{Co}}$, as a function of V thickness. There are two parameters that control the energy of the system, $\Delta\phi$ and the V thickness, and these parameters give rise to an energy landscape, as Fig. 1 shows. The minimum energy is indicated by a thick black line. The angle between the magnetization of the Fe and Co layers does not simply switch between 0° and 180°, as is the case for the conventional multilayers. Instead a more involved behaviour occurs, where, starting from a few atomic V layers, the coupling oscillates in a damped fashion, approaching 90° for thick V layers. This finding is a direct consequence of the competition between the MAE and the interlayer exchange coupling. Our result has consequences for the detection of small external fields, because the proposed multilayer, for low V thicknesses, is already more sensitive than a conventional device. For instance, at a V thickness of 5–6 atomic layers, a very small energy is needed to change the magnetization direction from about 130° to 90°. This can be seen in Fig. 1 as only two contours in the energy landscape must be passed. In a similar way, V layers of about 4 and 7 atomic widths give rise to angles of about 30° and 70°, respectively, which can be changed to 0° by a minute external field (ferromagnetic coupling). As the

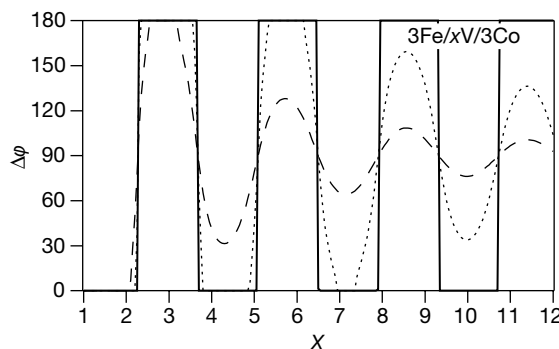


Figure 2 Calculated values of $\Delta\phi$, as a function of V thickness in atomic widths. The dashed curve is the same as the black curve in Fig. 1, from a calculation where the strength of the interlayer exchange coupling has been reduced by a factor of 5. The dotted

curve represents a calculation where no scaling has been made for the interlayer exchange coupling, and the solid curve represents a calculation where the magneto-crystalline anisotropy contribution has been neglected.

resistance of multilayers depends on the magnetic coupling, a small external field can be registered.

Finally, we note that although the basic idea behind this work is that new non-collinear magnetic configurations of multi-layer systems can be stabilized, the actual numbers calculated here for appropriate thicknesses are connected to some uncertainties. We have assumed, for instance, that the contribution given by the MAE in Fig. 1 is insensitive to the influence of the spacer layer. In addition, we have neglected the magnetic dipole contribution⁵ (which is indeed small for this system). The largest cause of error presented here is the calculated value of the interlayer exchange coupling, since previous work shows that the calculated strength is typically overestimated, compared to experimental measurements, by a factor of 5–10. In order to mimic an experimental situation as closely as possible we therefore reduced the calculated strength of the interlayer exchange coupling by a factor of 5. In order to analyse how different values of the interlayer exchange coupling influences the proposed multilayer we show in Fig. 2 the difference in angle, $\Delta\phi$, as a function of V thickness, for calculations with different strengths of the interlayer exchange coupling. The most conspicuous feature of the behaviour of $\Delta\phi$ as a function of spacer thickness is not modified by the strength of the exchange coupling; merely, the damping strength is influenced. Also, without the influence of the MAE the magnetization simply oscillates between ferro- and antiferromagnetic, whereas a non-negligible MAE produces a smoother behaviour in the coupling between different layers. Hence, these sources of error do not change the basic idea put forward here, that non-collinear, essentially perpendicular magnets may be grown, and that this is caused by a competition between interlayer exchange coupling and the MAE. This competition produces a magnetization profile that varies more smoothly with spacer thickness, and also reduces the energy between different magnetic configurations. We note here the difference from a conventional exchange-coupled sensor, that in theory could be tuned to a small coupling energy by getting close to a node in the oscillating exchange coupling curve. Owing to the large sensitivity to the exact thickness, such a set-up is useless in reality and has not been realized experimentally. In our proposed device, we suggest tuning the spacer-layer thickness to a local maximum in coupling strength; this is done in practice in conventional sensors.

The suggested multilayer sensor has, in certain aspects, a similar performance to conventional sensors, but with increased sensitivity, and from a technical process viewpoint we can draw on the experience of these systems. The precise number of atomic layers needed to optimize the proposed device must be determined experimentally. An extensive experimental study is needed to fine-tune the optimal choice of materials combinations and thicknesses. Other candidates of interest, in the sense that they have MAE stabilizing perpendicular magnetism in a strained state, are Ni and Co separated by an fcc spacer layer such as Cu, Pd or Pt. In addition, Pt has a large spin-orbit coupling that enhances the MAE values. □

Received 1 October 1999; accepted 17 May 2000.

1. Baibich, M. N. *et al.* Giant magnetoresistance of (001)Fe/(001)Cr magnetic superlattices. *Phys. Rev. Lett.* **61**, 2472–2475 (1988).
2. Parkin, S. S. P., More, N. & Roche, K. P. Oscillations in exchange coupling and magnetoresistance in metallic superlattice structures: Co/Ru, Co/Cr, and Fe/Cr. *Phys. Rev. Lett.* **64**, 2304–2307 (1990).
3. Binash, G., Grünberg, P., Saurenbach, F. & Zinn, W. Enhanced magnetoresistance in layered magnetic structures with antiferromagnetic interlayer exchange. *Phys. Rev. B* **39**, 4828–4830 (1989).
4. Bruno, P. Theory of interlayer magnetic coupling. *Phys. Rev. B* **52**, 411–439 (1995).
5. Aharoni, A. *Introduction to the Theory of Ferromagnetism* (Oxford Science, Oxford, 1996).

Acknowledgements

Support from J. M. Wills is acknowledged. This project has been financed by the Swedish Natural Science and Technical Research Councils (NFR and TFR). Support for the European programme, Training and Mobility of Researchers (TMR) is acknowledged.

Correspondence and requests for materials should be addressed to B.J. (e-mail: borje.johansson@fysik.uu.se).

Imaging the vortex-lattice melting process in the presence of disorder

Alex Soibel*, Eli Zeldov*, Michael Rappaport†, Yuri Myasoedov*, Tsuyoshi Tamegai‡, Shuichi Ooi ‡, Marcin Konczykowski§ & Vadim B. Geshkenbein||

* Department of Condensed Matter Physics, The Weizmann Institute of Science, Rehovot 76100, Israel

† Physics Services, The Weizmann Institute of Science, Rehovot 76100, Israel

‡ Department of Applied Physics, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

§ CNRS, UMR 7642, Laboratoire des Solides Irradiés, Ecole Polytechnique, 91128 Palaiseau, France

|| Theoretische Physik, ETH-Honggerberg, CH-8093 Zurich, Switzerland,

& L. D. Landau Institute for Theoretical Physics, 117940 Moscow, Russia

¶ CREST, Japan Science and Technology Corporation (JST), Japan

General arguments¹ suggest that first-order phase transitions become less sharp in the presence of weak disorder, while extensive disorder can transform them into second-order transitions; but the atomic level details of this process are not clear. The vortex lattice in superconductors provides a unique system in which to study the first-order transition^{2–6} on an inter-particle scale, as well as over a wide range of particle densities. Here we use a differential magneto-optical technique to obtain direct experimental visualization of the melting process in a disordered superconductor. The images reveal complex behaviour in nucleation, pattern formation, and solid–liquid interface coarsening and pinning. Although the local melting is found to be first-order, a global rounding of the transition is observed; this results from a disorder-induced broad distribution of local melting temperatures, at scales down to the mesoscopic level. We also resolve local hysteretic supercooling of microscopic liquid domains, a non-equilibrium process that occurs only at selected sites where the disorder-modified melting temperature has a local maximum. By revealing the nucleation process, we are able to experimentally evaluate the solid–liquid surface tension, which we find to be extremely small.

We first discuss the expected vortex-lattice melting process in the absence of disorder. Under equilibrium magnetization conditions in platelet-shaped samples in perpendicular applied field $H_a \parallel z$, the internal fields $B(x, y)$ and $H(x, y)$ across the sample have a dome-shaped profile with a maximum at the centre⁷. This is because in the absence of bulk pinning, the equilibrium shielding currents flow only along the sample edges. As H_a or temperature T is increased, the field H in the central part of the sample reaches the melting field $H_m(T)$ first, and thus a small round ‘puddle’ of vortex liquid should be formed in the centre, surrounded by vortex solid. Because of the first-order nature of the transition, the vortex-lattice melting is associated with a discontinuous step in the equilibrium magnetization⁶, $4\pi\Delta M = \Delta(B - H)$. Because in our geometry H is continuous across the solid–liquid interface, the field B in the liquid is enhanced by ΔB relative to the solid. In $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_8$ (BSCCO) crystals ΔB is typically⁶ 0.1–0.4 G. Conventional magneto-optical (MO) imaging techniques^{8–10} (Fig. 1 legend) cannot resolve such small field differences. We have therefore devised the following differential method.

An MO image is acquired by averaging typically ten charge-coupled device (CCD) images at some H_a and T . Then H_a is increased by $\delta H_a \ll H_a$, or T is increased by $\delta T \ll T$, a second averaged image is obtained, and subtracted from the first. This process is averaged typically 100 times, yielding a differential field resolution of about 30 mG, approximately two orders of magnitude better than the standard MO method. By recording the differential

images in a sequence of fields or temperatures a ‘movie’ of the melting process is obtained.

As H_a is increased by δH_a , the radius of the vortex-liquid puddle should ideally increase by δR , determined by the gradient of the dome-shaped profile $H(x,y)$. The differential image in this case should show a bright ring on a dark background, which indicates the location of the expanding solid–liquid interface. In the rest of the image almost no change in the field should occur, except the uniform background signal of δH_a . The intensity of the ring is ΔB above the background¹¹, independent of δH_a , whereas the width of the ring reflects the distance δR over which the interface expands due to δH_a . This is the expected equilibrium coexistence of the solid and liquid phases due to an intrinsic gradient of the internal field, similar to the phase separation of water and ice in the presence of a gravitational field. However, we find that the phase separation is governed by the disorder, and the dome-shaped field profile merely delays the melting process near the edges.

Figure 1a presents differential MO images from the vortex-lattice melting ‘movie’ in one of the smaller BSCCO crystals, which is initially in the vortex solid phase. At 159.5 Oe a small liquid puddle is nucleated, seen as a bright spot in the upper-right corner. In contrast to expectations, the puddle is not in the centre of the sample, nor is it round; instead, a rather rough shape of the liquid domain is observed. As H_a is further increased a ring-like bright object is obtained, which is the solid–liquid interface separating the liquid from the surrounding vortex solid. Both the shape and the

width of the ring are highly non-uniform. At 165 Oe a ‘tongue’ of the liquid protrudes sharply to the left side. Figure 1b shows the corresponding outer contours of the liquid phase at 0.5 Oe intervals of H_a . The interface is often pinned, resulting in overlapping contours, and then bursts out to more remote locations by a sequence of irregular local protrusions with a high degree of corrugation.

We now show that the observed complexity of the melting process is the result of disorder. Theoretical arguments¹ suggest that disorder in solids causes local variations in the melting field H_m , resulting in a rough $T_m(x,y)$ or $H_m(x,y)$ ‘landscape’, and the distribution function of the local melting field $f_m(H)$ develops characteristic tails on both sides of the mean field H_m^0 . In high-temperature superconductors, point disorder as well as anisotropy are known to shift¹² H_m^0 , and therefore local variations in these parameters¹³, quenched during crystal growth, should be the main source of the roughening of the $H_m(x,y)$ landscape and the broadening of $f_m(H)$, as illustrated in Fig. 2a.

In order to analyse this process quantitatively, Fig. 2b shows a partial set of melting images in a large BSCCO crystal with three visible crystallographic defects. The vortex-liquid phase of a very irregular shape is formed initially at 91.5 Oe in the lower left corner (blue). At 92.5 Oe a separate liquid region appears along one of the defects. With increasing H_a very complex patterns of intermixed solid and liquid regions are formed. The dominant effect of the crystallographic disorder is clearly visible here because the melting

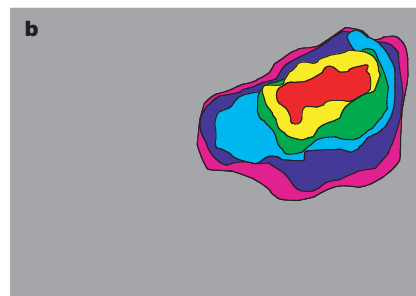
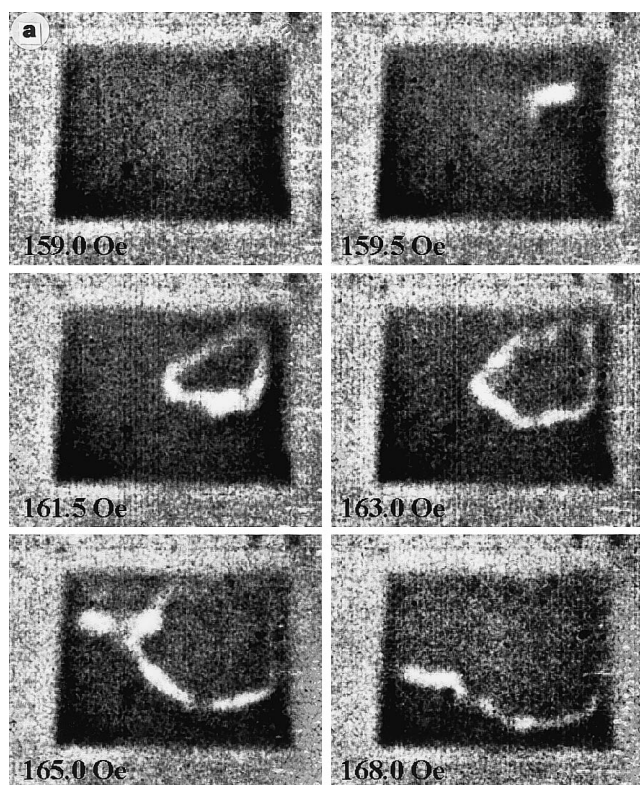


Figure 1 Vortex-lattice melting process in a small BSCCO crystal. Magneto-optical (MO) imaging^{8–10} is achieved by placing a garnet indicator film on a sample. Linearly polarized light undergoes Faraday rotation in the indicator and is reflected back through a crossed polarizer, resulting in a real-time image of the magnetic field distribution.

a, Differential MO images of the melting process in a BSCCO crystal ($T_c = 60$ K) of area 0.35×0.27 mm² at $T = 60$ K and $H_a \parallel c$ -axis. (The full movie is available at <http://www.weizmann.ac.il/home/fnsup>). The grayscale from black to white spans a field range of 0.2 G. The region outside the sample is bright. The differential images are obtained by subtracting the image at field H_a from the image at $H_a + \delta H_a$, with $\delta H_a = 1$ Oe. At 159.5 Oe a small liquid droplet is nucleated within a solid resulting in the bright spot. Increasing H_a to 161.5 and 163.0 Oe, the liquid phase expands, seen as a dark

region surrounded by a bright ring. The local width of the ring reflects the extent of the propagation of the solid–liquid interface due to the field modulation δH_a . The entire upper part is in the liquid phase at 168.0 Oe. **b**, Contours of the liquid phase at $H_a = 159.5, 160.0, 160.5, 161.0, 161.5$ and 162.0 Oe. The propagation of the solid–liquid interface is characterized by pinning and local protrusions resulting in a rough structure. The contours can be viewed as equipotential lines of $H_m(x,y)$ landscape (see Fig. 2a). Overlapping contours indicate an interface which is pinned due to steep variations in $H_m(x,y)$. As H_a is increased from 159.5 Oe (red) to 160.0 Oe (yellow) the interface remains pinned at the right and left edges, but expands upwards and downwards. The upper part remains pinned at 160.5 Oe (green) and 161.0 Oe (cyan), then expands abruptly at 161.5 Oe and gets pinned again at 162.0 Oe.

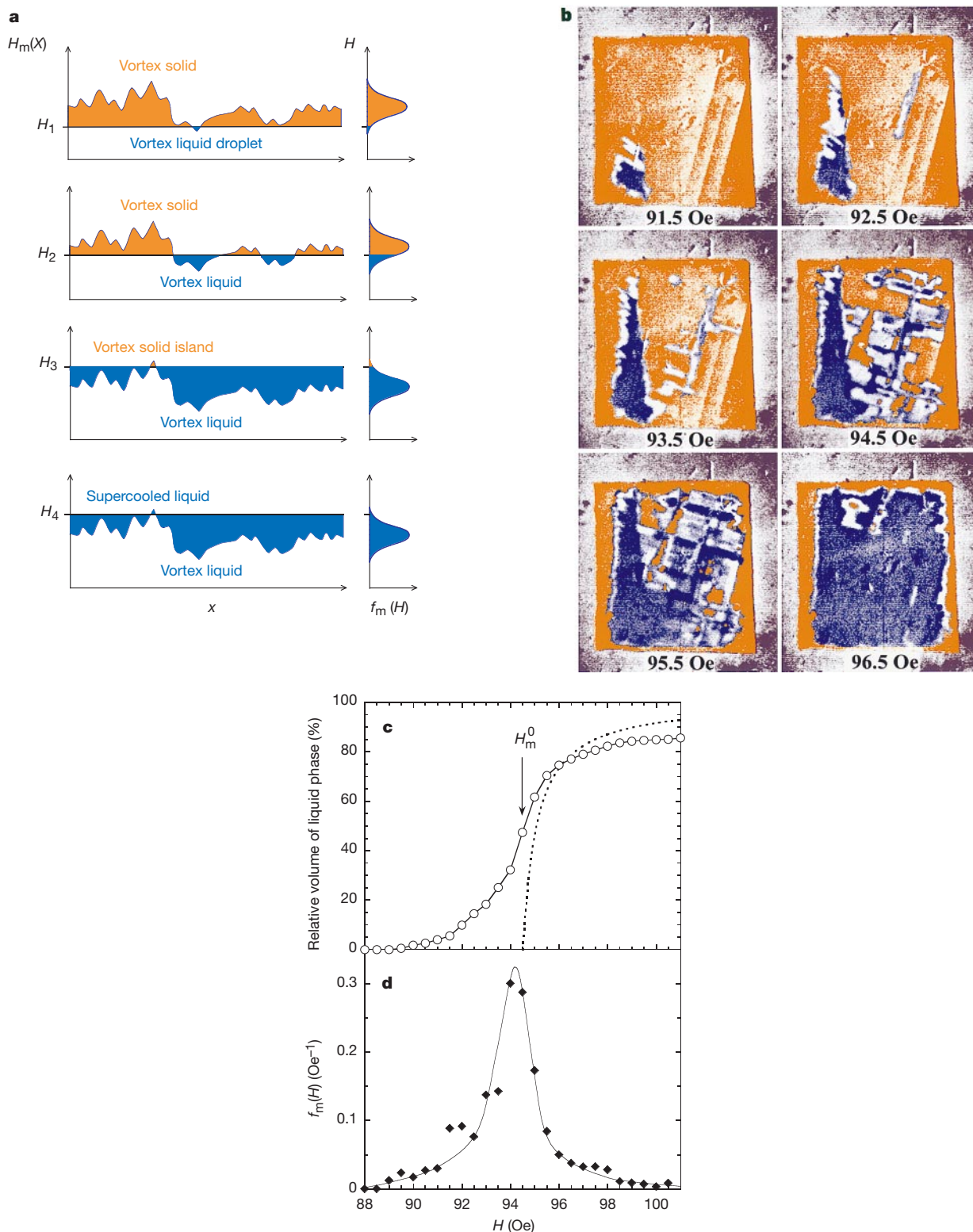


Figure 2 Melting process in the presence of disorder. **a**, Schematic plot of $H_m(x)$ landscape and the distribution function $f_m(H)$, for various values of the field H . When $H = H_1$ in the low-field tail of $f_m(H)$, a small liquid droplet (blue) is nucleated at the minimum of $H_m(x)$. As H is increased to H_2 , large liquid domains are formed. A small solid island is present within the liquid at H_3 , in the high-field tail of $f_m(H)$. After complete melting and upon reduction of the field to $H_4 = H_3$, the solid island is found to be absent, giving rise to a supercooled liquid domain and hysteretic local behaviour. **b**, Melting process in a large BSCCO crystal of $1.1 \times 1.2 \times 0.025 \text{ mm}^3$ at $T = 70 \text{ K}$ and $\delta H_a = 1 \text{ Oe}$, demonstrating the effects of disorder. (The full movie is available at <http://www.weizmann.ac.il/home/fnsup>). The solid and liquid regions are shown in brown and blue, respectively. At 91.5 Oe an irregular liquid droplet is formed in the lower left corner. On the right-hand side three parallel crystallographic defects

are visible. The crystallographic disorder in the sample has characteristic directions aligned parallel and perpendicular to the crystal growth direction, inducing corresponding features in the $H_m(x, y)$ landscape. As a result, complicated melting patterns are resolved at 93.5, 94.5 and 95.5 Oe with numerous solid and liquid domains. At 96.5 Oe a large liquid region is present in the centre, with a few solid islands in the top part. **c**, Liquid volume fraction versus H_a obtained from the full melting sequence, showing the global rounding of the melting transition with characteristic tails above and below the mean field H_m^0 . The dashed line shows the calculated liquid fraction in the absence of disorder, due to the dome-shaped internal field profile. **d**, The experimental distribution function of the melting field $f_m(H)$ obtained by taking the derivative of the vortex liquid volume in **c** with respect to the field H . The solid line is a guide to the eye.

patterns have well defined preferential directions aligned parallel and perpendicular to the defects. Figure 2c presents the volume fraction of the liquid versus H_a , which allows a derivation of the distribution function $f_m(H)$ shown in Fig. 2d. Although locally the melting is a sharp first-order transition, the global solid–liquid transformation is completely rounded, and $f_m(H)$ displays pronounced tails extending both below and above H_m^0 . This general form of $f_m(H)$, which we observe in all the samples, is in agreement with the theoretical predictions¹. The dashed line in Fig. 2c shows for comparison the calculated melting behaviour in the absence of disorder, resulting from the fit to the experimental dome-shaped field profile across the sample. In this case the melting should set in abruptly, and most of the volume should melt within a few Oe from H_m^0 . Thus the disorder is dominant in the observed rounding of the transition and in the development of $f_m(H)$ tails.

We present several examples of the melting features in various crystals. Figure 3a shows an enlarged view of nucleation of two liquid droplets in the low-field tail of $f_m(H)$. The width of the

smaller droplet is below $6\text{ }\mu\text{m}$ (close to our resolution limit), which means that this droplet is less than seven vortices wide. This image is a striking demonstration that the disorder-nucleated melting occurs on a mesoscopic level. Furthermore, it shows that the solid–liquid surface tension is very low, as addressed below. Such liquid droplets, formed at local minima of $H_m(x,y)$ (see Fig. 2a), are found to behave reversibly with field, in contrast to solid islands at maxima of $H_m(x,y)$, as described below. Another example of liquid nucleation in a different BSCCO crystal is shown in Fig. 3b. Here the nucleation occurs simultaneously in three adjacent regions, demonstrating specific correlations in the disorder distribution quenched during crystal growth. The characteristic width of the solid and liquid patterns is about $25\text{ }\mu\text{m}$ (34 vortices), further indicating low surface tension. High surface tension would prevent such proximity of microscopic domains, favouring formation of a single larger liquid region. With increasing H_a large liquid domains of irregular shape are formed, as demonstrated in Fig. 3c. Here a highly irregular ring separates the vortex liquid from the surrounding solid. The local

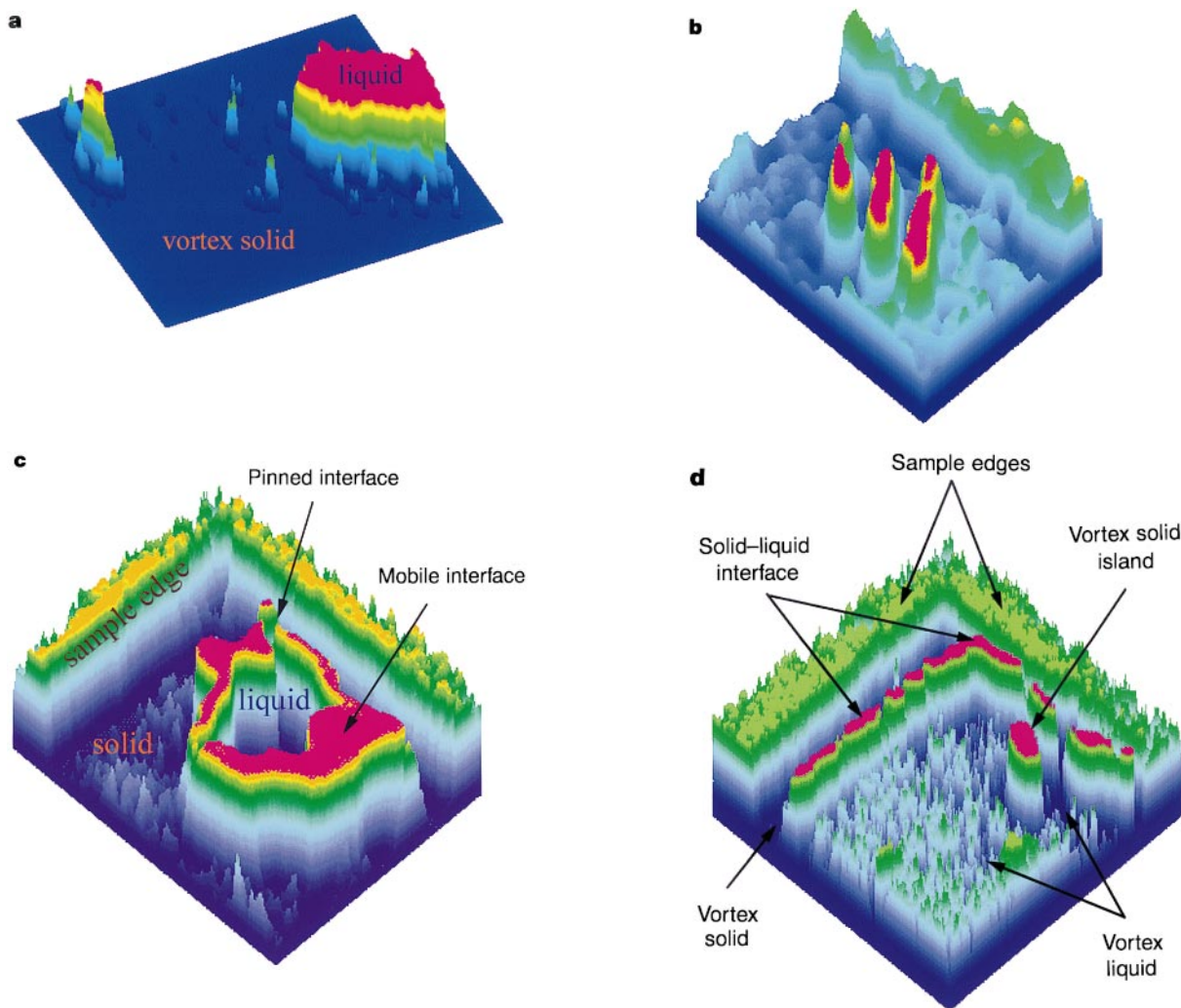


Figure 3 Four examples of melting features in various BSCCO crystals shown in a three-dimensional representation. The field scale from blue to magenta is about 0.3 G. Only selected parts of the samples are shown for clarity. **a**, Two liquid droplets (magenta) surrounded by vortex solid (blue). The smaller droplet on the left is only about seven vortices wide, demonstrating that disorder-assisted melting occurs on the scale of the inter-vortex distance. Image area $0.07 \times 0.14\text{ mm}^2$, $H_a = 25.75\text{ Oe}$, $\delta H_a = 0.5\text{ Oe}$, and $T = 84\text{ K}$. **b**, Vortex liquid nucleation (magenta) occurring simultaneously in three adjacent finger-like regions. The image size is $0.24 \times 0.31\text{ mm}^2$, $H_a = 38.5\text{ Oe}$, $\delta H_a = 1\text{ Oe}$, and $T = 80\text{ K}$. The top right side of the image shows one of the sample edges (green). **c**, An irregular vortex liquid region surrounded by vortex solid. The solid–liquid interface (magenta) is highly corrugated. In some regions the interface appears to be very

wide due to its high mobility during the field modulation by δH_a . At other points the interface appears to be discontinuous due to interface pinning, which results in a lack of differential signal during field modulation. Outer edges of the sample are seen along the top sides of the image. $H_a = 83\text{ Oe}$, $\delta H_a = 0.5\text{ Oe}$, and $T = 70\text{ K}$. Image area is $0.26 \times 0.34\text{ mm}^2$. **d**, A small vortex-solid island surrounded by vortex liquid. The liquid phase is in turn surrounded by vortex solid along the sample edges. The solid–liquid interface near the edges (magenta rim) appears highly non-uniform and discontinuous due to the alternating regions of mobile and pinned interface. The melting and freezing process of this solid island is characterized by a high degree of hysteresis, as shown in Fig. 4. Image area $0.68 \times 0.68\text{ mm}^2$, $H_a = 69.5\text{ Oe}$, $\delta H_a = 0.5\text{ Oe}$, and $T = 75\text{ K}$.

width of the ring is a direct measure of the distance the interface expands due to δH_a . In some parts the ring is very broad, showing the locations where $H_m(x,y)$ has smooth, gradual behaviour (see Fig. 2a), resulting in a very mobile interface. In contrast, in other places where the ring appears to have discontinuities, the solid–liquid interface is effectively pinned due to steep $H_m(x,y)$. As the liquid domains expand with H_a , the peaks in $H_m(x,y)$ melt last, forming solid islands within the liquid (see Fig. 2a). Figure 3d depicts such a small solid island completely surrounded by vortex liquid, which in turn is surrounded by vortex solid along the edges of the crystal. In this configuration non-equilibrium phenomena are found to be present, as described below.

Our experimental results allow evaluation of the surface tension σ of the solid–liquid interface. Theoretically, the surface tension can be estimated as follows. At the first-order transition the discontinuous change in the internal energy is compensated by a change in the entropy, such that the free energy remains continuous. The jump in the energy density is roughly equal to $H_m \Delta B / 4\pi$. If the solid–liquid interface has a width of the order of the inter-vortex spacing a_0 , then within this width the system is neither liquid nor solid. Thus the energy cost per unit area of this interface can be estimated as $\sigma = \eta a_0 H_m \Delta B / 4\pi$, where η is an unknown numerical factor presumed to be of order unity. When H exceeds a local minimum by $h = H - H_m$, a liquid droplet should be formed in a solid, provided the energy gain in melting is larger or equal to the energy cost of forming an interface. Assuming that the droplet is of radius r and that it extends throughout the thickness d of the sample, the energy gain is given approximately by $(\partial F_s / \partial H - \partial F_l / \partial H) h \pi r^2 d = \Delta B h r^2 d / 4$. Here $\partial F_s / \partial H$ and $\partial F_l / \partial H$ are

the derivatives of the free energies of the solid and liquid phases respectively, and we have used the fact that $\partial F_s / \partial H - \partial F_l / \partial H = \Delta B / 4\pi$ at the melting transition. By comparing this gain with the interface energy $2\pi \sigma r d$, we obtain the minimum nucleation radius $r_{\min} \approx 2\eta a_0 H_m / h$. The presented images show the existence of very small liquid droplets with $r_{\min} \approx 3a_0$. By analysing the expansion of the droplets due to δH_a and upon varying H_a , we estimate the superheating h to be less than 0.5 Oe for a typical H_m of 50 Oe, which results in $\eta \leq 10^{-2}$. Thus, assuming that $H_m(x,y)$ exhibits smooth local behaviour, our measurements indicate that σ is two orders of magnitude lower than rough theoretical estimates. Such a low surface tension implies that the vortex melting process is completely governed by disorder and that the stabilizing influence of the interface¹ can be neglected. The origin of this phenomenon requires further investigation.

Hysteretic supercooling and superheating is an inherent property of a first-order transition¹⁴, but the precise mechanism on a microscopic level is not clear. We are able to spatially resolve this unique feature, as shown in Fig. 4. We find that the hysteresis occurs only at the local maxima of $H_m(x,y)$, indicating that disorder is an important factor also in the non-equilibrium properties. In the top left image, at 68.50 Oe, a few vortex–solid islands are visible within the liquid in the upper part. As the field is increased, these solid islands split into smaller ones, and gradually shrink and disappear. At 69.50 Oe only one small island is left, visible as a bright spot, which fully melts at 69.75 Oe. A better view of the top-left corner of the sample including this island is shown in Fig. 3d. Upon decreasing H_a , the bright interface along the sample edges that separates the large central liquid region and the outer solid strips behaves

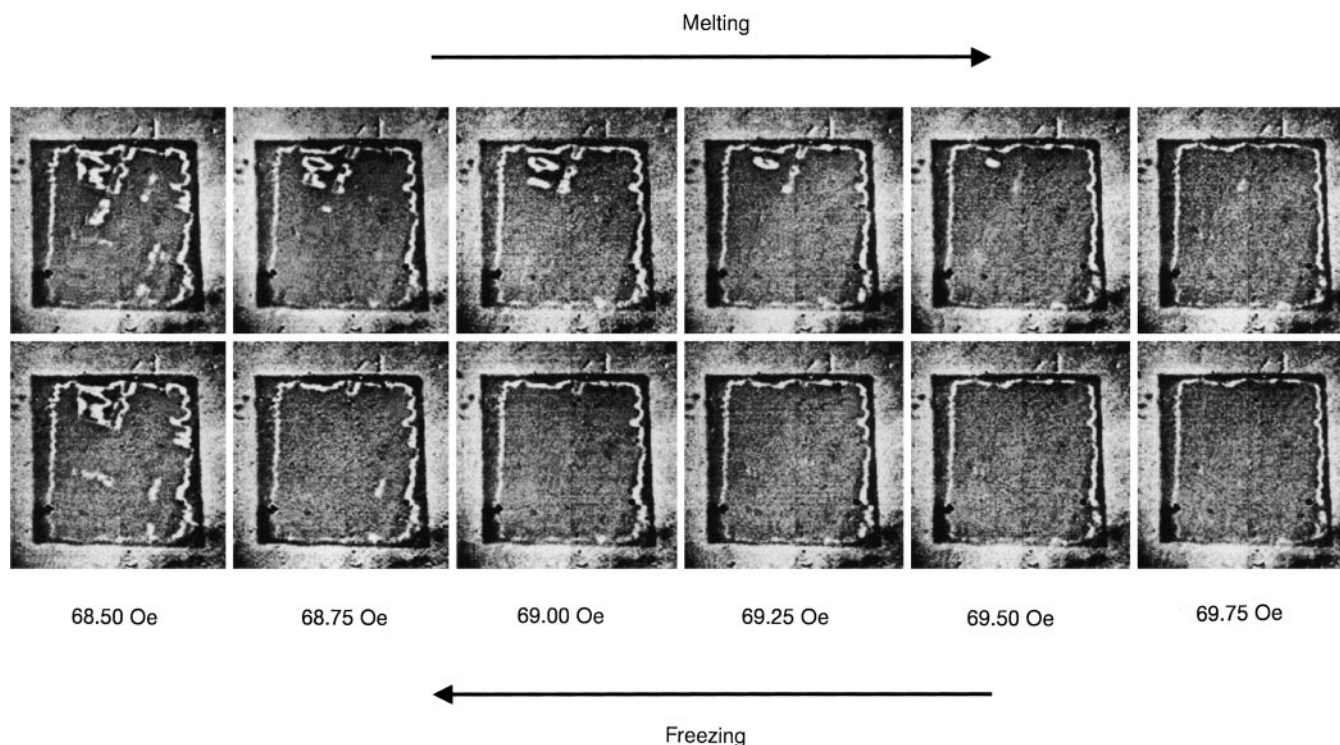


Figure 4 Direct visualization of the local supercooling and hysteresis process. The top row of images shows the melting sequence upon increasing the field, while the bottom shows the freezing process at the same values of decreasing field. The central region of the sample is in the liquid phase surrounded by vortex solid along the edges because of the dome-shaped field profile. The bright solid–liquid interface near the edges is highly corrugated due to the disorder-induced variations in the local $H_m(x,y)$, but it is completely reversible upon melting and freezing. In contrast, the behaviour of the solid islands, visible in the upper left part of the sample within the large liquid region, is highly hysteretic. In the melting sequence the islands shrink gradually until the last of the islands fully melts at 69.75 Oe. In the freezing sequence, however, the islands are absent and suddenly

reappear only at 68.50 Oe, while the elongated central island nucleates at still lower field (not shown). In the four freezing images, 68.75 Oe through 69.50 Oe, in the areas where the islands are absent relative to the corresponding melting images, the liquid phase is in a supercooled state. These metastable supercooled liquid regions coexist with the surrounding equilibrium liquid. Such local supercooling is found only at local maxima of $H_m(x,y)$. No corresponding superheating of the solid is found at the minima of $H_m(x,y)$, and no hysteresis is observed in the regions where $H_m(x,y)$ varies monotonically. Sample size is $1.1 \times 1.2 \times 0.025 \text{ mm}^3$, $T = 75 \text{ K}$ and $\delta H_a = 0.5 \text{ Oe}$. Small dark and bright spots that do not vary with field are defects in the MO indicator.

completely reversibly (compare the melting and freezing images at the same field). The inner solid islands, in contrast, are absent upon freezing, and then suddenly reappear at a lower field of 68.50 Oe. Once formed, the islands extend to their full size at the corresponding field, and proceed to behave reversibly as long as they are not fully melted again.

Reversible behaviour most probably reflects an equilibrium state of the system. The above results therefore imply that there is no supercooling or superheating when the solid–liquid interface is located in a region where the $H_m(x,y)$ landscape varies monotonically. The hysteretic behaviour occurs only when a new interface has to be formed, but only if this new interface surrounds a solid rather than a liquid. Hence, upon reducing H to below a local maximum of $H_m(x,y)$, the solid island is not readily formed (see lower panel of Fig. 2a). In this regime a unique situation occurs: metastable supercooled liquid domains, which are present instead of the ‘missing’ solid islands, coexist with the surrounding equilibrium liquid. Only when the supercooling becomes sufficiently large does sudden nucleation occur, upon which the solid abruptly occupies the entire volume of the supercooled domains. This non-equilibrium process is observed only at local maxima of $H_m(x,y)$. We would expect a symmetric non-equilibrium mechanism to be present at minima of $H_m(x,y)$, where superheated solid would transform hysteretically into a liquid droplet, but such hysteresis is not found. The asymmetry between the supercooling and superheating on a macroscopic scale is usually ascribed to the surface wetting process, which prevents solid superheating¹⁴. In our case mesoscopic vortex droplets are formed within a solid and therefore such an asymmetry should not be expected. Nevertheless, the tips of the vortices near the top and bottom surfaces of the crystal experience a reduced elastic confinement potential due to the absence of the lattice outside the sample. As a result, the vortex tips may undergo a pre-melting transition similar to the surface wetting in atomic solids, which may prevent the superheating of the vortex solid. In contrast to atoms, however, each vortex tip is attached to a solid vortex in the bulk, and therefore is restricted in space and may not exhibit liquid properties. □

Received 10 April; accepted 27 May 2000.

- Imry, Y. & Wortis, M. Influence of quenched impurities on first-order phase transitions. *Phys. Rev. B* **19**, 3580–3585 (1980).
- Blatter, G. *et al.* Vortices in high-temperature superconductors. *Rev. Mod. Phys.* **66**, 1125–1388 (1994).
- Nelson, D. R. Vortex entanglement in high- T_c superconductors. *Phys. Rev. Lett.* **60**, 1973–1976 (1988).
- Safar, H. *et al.* Experimental evidence for a first-order vortex-lattice-melting transition in untwinned single crystal $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$. *Phys. Rev. Lett.* **69**, 824–827 (1992).
- Kwok, W. K. *et al.* Vortex lattice melting in untwinned and twinned single-crystals of $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$. *Phys. Rev. Lett.* **69**, 3370–3373 (1992).
- Zeldov, E. *et al.* Thermodynamic observation of first-order vortex-lattice melting transition. *Nature* **375**, 373–376 (1995).
- Zeldov, E. *et al.* Geometrical barriers in high-temperature superconductors. *Phys. Rev. Lett.* **73**, 1428–1431 (1994).
- Indenbom, M. V. *et al.* Direct study of magnetic flux penetration and trapping in HTSC. *Physica C* **166**, 486–496 (1990).
- Duran, C. A. *et al.* Real-time imaging of the magnetic flux distribution in superconducting $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$. *Nature* **357**, 474–477 (1992).
- Welp, U. *et al.* Imaging of transport currents in superconducting $(\text{Bi, Pb})_2\text{SrCa}_2\text{Cu}_3\text{O}_x$ composites. *Nature* **376**, 44–46 (1995).
- Morozov, N. *et al.* Paramagnetic ac susceptibility at first-order vortex-lattice phase transition. *Phys. Rev. B* **54**, R3784–R3787 (1996).
- Khaykovich, B. *et al.* Vortex-matter phase transitions in $\text{Bi}_2\text{Sr}_2\text{CaCu}_3\text{O}_{8-x}$: Effects of weak disorder. *Phys. Rev. B* **56**, R517–R520 (1997).
- Oral, A. *et al.* Disorder-driven intermediate state in the lattice melting transition of $\text{Bi}_2\text{Sr}_2\text{CaCu}_3\text{O}_{8-x}$ single crystals. *Phys. Rev. B* **56**, R14295–R14298 (1997).
- Landau, L. D. & Lifshitz, E. M. *Statistical Physics* 3rd edn (Pergamon, Oxford, 1993).

Acknowledgements

We thank D. R. Nelson, Y. Imry, V. Vinokur, L. Balents, C. Varma, C. van der Beek and M. Indenbom for valuable discussions. This work was supported by the Israel Science Foundation and Center of Excellence Program, by the Ministry of Science, Israel, by the MINERVA Foundation, Munich, Germany, and by the Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture, Japan.

Correspondence and requests for materials should be addressed to A. S. (email: hsasha@wisemail.weizmann.ac.il).

Optimal shapes of compact strings

Amos Maritan*, Cristian Micheletti*, Antonio Trovato* & Jayanth R. Banavar†

* International School for Advanced Studies (SISSA), Via Beirut 2–4, 34014 Trieste, Istituto Nazionale per la Fisica della Materia (INFM) and the Abdus Salam International Center for Theoretical Physics, Trieste, Italy

† Department of Physics and Center for Materials Physics, 104 Davey Laboratory, The Pennsylvania State University, University Park, Pennsylvania 16802, USA

Optimal geometrical arrangements, such as the stacking of atoms, are of relevance in diverse disciplines^{1–5}. A classic problem is the determination of the optimal arrangement of spheres in three dimensions in order to achieve the highest packing fraction; only recently has it been proved^{1,2} that the answer for infinite systems is a face-centred-cubic lattice. This simply stated problem has had a profound impact in many areas^{3–5}, ranging from the crystallization and melting of atomic systems, to optimal packing of objects and the sub-division of space. Here we study an analogous problem—that of determining the optimal shapes of closely packed compact strings. This problem is a mathematical idealization of situations commonly encountered in biology, chemistry and physics, involving the optimal structure of folded polymeric chains. We find that, in cases where boundary effects⁶ are not dominant, helices with a particular pitch-radius ratio are selected. Interestingly, the same geometry is observed in helices in naturally occurring proteins.

The problem of placing spheres in three dimensions in order to attain the highest density was first posed by Kepler and has attracted much interest, culminating in its recent rigorous mathematical solution¹. The close-packed hard-sphere problem may be restated in an alternative manner, more convenient for numerical implementation, as the determination of the arrangement of a set of points in a given volume that results in the minimum distance between any pair of points being as large as possible⁶. It is notable that the resulting ‘bulk’ optimal arrangement exhibits translational invariance in that, far from the boundaries, the local environment is the same for all points.

Here we introduce a new problem pertaining to the optimal shapes of compact strings. We consider a string (an open curve) in three dimensions. We use a geometric measure⁷ of the curve, the ‘rope length’, defined as the arc length measured in units of the thickness, which has proved to be valuable in applications of knot theory^{7–12}. The thickness Δ denotes the maximum radius of a uniform tube with the string passing through its axis, beyond which the tube either ceases to be smooth, owing to tight local bends, or it self-intersects. Our focus is on finding the optimal shape of a string of fixed arc length, subject to constraints of compactness, which would maximize its thickness, or equivalently minimize its rope length.

Following the approach of Gonzalez and Maddocks¹⁰, who studied knotted strings, we define a global radius of curvature as follows. The global radius of curvature of the string at a given point is computed as the minimum radius of the circles going through that point and all other pairs of points of the string. It generalizes the concept of the local radius of curvature (the radius of the circle which locally best approximates the string) by taking into account both local (bending of the string) and non-local (proximity of another part of the string) effects. For discretized strings the local radius of curvature at a point is simply the radius of the circle going through the point and its two adjoining points. The minimum of all the global radii then defines the thickness, that is, the minimum radius of the circles going through any triplet of discrete points. This coincides with the previous definition in the continuum limit,

obtained on increasing the number of discretized points (assumed to be equally spaced) on the string while keeping the string length fixed¹⁰.

We used several different boundary conditions to enforce the confinement of the string. The simplest ones discussed here are the confinement of a string of length l within a cube of side L or constraining it to have a radius of gyration (which is the r.m.s. distance of the discretized points from their centre of mass) that is less than a pre-assigned value R . Even though different boundary conditions influence the optimal string shape, the overall features are found to be robust. Examples of optimal shapes, obtained from numerical simulations, for different ratios of l/L and l/R are shown in Fig. 1. In both cases, two distinct families of strings, helices and saddles, appear. The two families are close competitors for optimality and different boundary conditions may stabilize one over the other. For example, if optimal strings of fixed length are constrained to have a radius of gyration less than R , then upon decreasing R , the string goes from a regime where the trivial linear string is curled into an arc, then into a portion of helix and finally into a saddle. When the string is constrained to lie within a cube of size L , as L decreases first saddles are observed and then helices.

We have also been able to find bulk-like solutions which are not influenced by boundary effects. Such solutions can be obtained by imposing uniform local constraints along the string. On imposing a minimum local density on successive segments of the string (for example, constraining each set of six consecutive beads to have a radius of gyration less than a preassigned value) we obtained perfectly helical strings, as in Fig. 2, confirming that this is the optimal arrangement. Note that, in close analogy with the sphere-packing problem, the optimal shape displays translational invariance along the chain. In all cases, the geometry of the chosen helix is such that there is an equality of the local radius of curvature

(determined by the local bending of the string) and the radius that is associated with a suitable triplet of non-consecutive points lying in two successive turns of the helix. This is a feature that is observed only for a special ratio c^* of the pitch, p , and the radius, r , of the circle projected by the helix on a plane perpendicular to its axis. When $p/r > c^* = 2.512$ the global radius of curvature is equal to the local radius with the helix thickness given by $\Delta = r(1 + p^2/(2\pi r)^2)$. If $p/r < c^*$, the global radius of curvature is strictly lower than the local radius, and the helix thickness is determined basically by the distance between two consecutive helix turns: $\Delta \approx p/2$ if $p/r \ll 1$. Optimal packing selects the very special helices corresponding to the transition between the two regimes described above. A visual example is provided by the optimal helix of Fig. 2.

For discrete strings, the critical ratio p/r depends on the discretization level. A more robust quantity is the ratio f , averaged over all the points of the string, of the minimum radius of the circles going through each point and any two non-adjacent points and the local radius. For discretized strings, $f = 1$ just at the transition described above, whereas $f > 1$ in the 'local' regime and $f < 1$ in the 'non-local' regime. In our computer-generated optimal strings, f differed from unity by less than one part in a thousand.

It is interesting to note that, in nature, there are many instances of the appearance of helices. For example, many biopolymers, such as proteins and enzymes, have backbones which frequently form helical motifs. (Rose and Seltzer¹³ have used the local radii of curvature of the backbone as an input in an algorithm for finding the peptide chain turns in a globular protein.) It has been shown¹⁴ that the emergence of such motifs in proteins (unlike in random heteropolymers which, in the melt, have structures conforming to gaussian statistics) is the result of the evolutionary pressure exerted by nature in the selection of native state structures that are able to

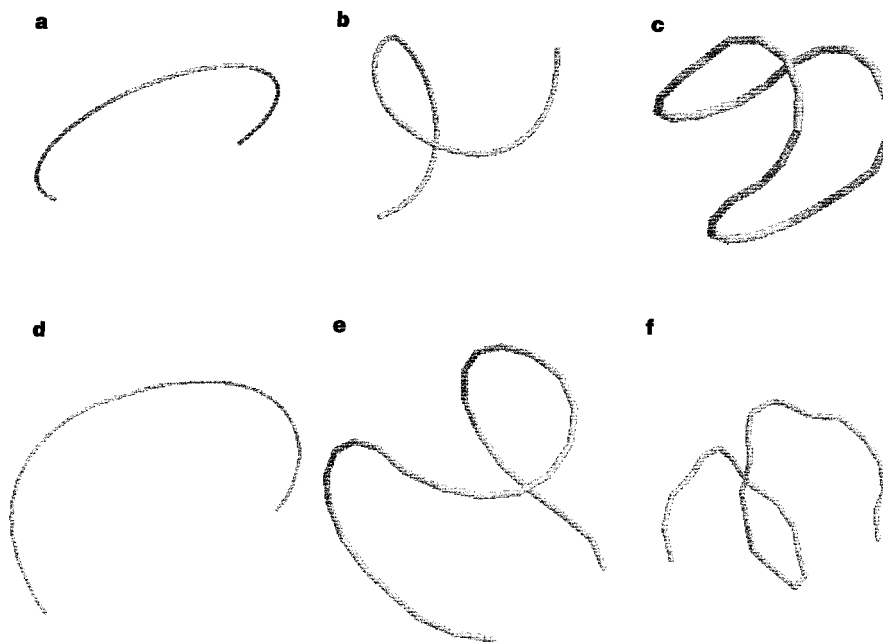


Figure 1 Examples of optimal strings. The strings in the figure were obtained starting from a random conformation of a chain made up of N equally spaced points (the spacing between neighbouring points is defined to be 1 unit) and successively distorting the chain with pivot, crankshaft and slithering moves commonly used in stochastic chain dynamics²⁰. A Metropolis Monte Carlo procedure is employed with a thermal weight, $e^{+\Delta/T}$, where Δ is the thickness and T is a fictitious temperature set initially to a high value such that the acceptance rate is close to 1 and then decreased gradually to zero in several thousand steps. Self-avoidance of the optimal string is a natural consequence of the maximization of the thickness. The introduction of a hard-core repulsion between beads

was found to significantly speed up convergence to the optimal solution and avoid trapping in self-intersecting structures. We have verified that the same values (within 1%) of the final thickness of the optimal strings are obtained starting from unrelated initial conformations. **a–c**, Optimal shapes obtained by constraining strings of 30 points with a radius of gyration less than R . **a**, $R = 6.0$, $\Delta = 6.42$. **b**, $R = 4.5$, $\Delta = 3.82$. **c**, $R = 3.0$, $\Delta = 1.93$. **d–f**, Optimal shapes obtained by confining a string of 30 points within a cube of side L . **d**, $L = 22.0$, $\Delta = 6.11$. **e**, $L = 9.5$, $\Delta = 2.3$. **f**, $L = 8.1$, $\Delta = 1.75$.

house sequences of amino acids which fold reproducibly and rapidly¹⁵ and are characterized by a high degree of thermodynamic stability¹⁶. Furthermore, because of the interaction of the amino acids with the solvent, globular proteins attain compact shapes in their folded states.

It is then natural to measure the shape of these helices and assess if they are optimal in the sense described here. The measure of f in α -helices found in naturally occurring proteins yields an average value for f of 1.03 ± 0.01 , hinting that, despite the complex atomic chemistry associated with the hydrogen bond and the covalent bonds along the backbone, helices in proteins satisfy optimal packing constraints (for the measure of f we considered α -helices extracted from the unrelated proteins 1env, 1beo and 2end in the Protein Data Bank). This result implies that the backbone sites in

protein helices have an associated free volume distributed more uniformly than in any other conformation with the same density. This is consistent with the observation¹⁴ that secondary structures in natural proteins have a much larger configurational entropy than other compact conformations. This uniformity in the free volume distribution seems to be an essential feature because the requirement of a maximum packing of backbone sites by itself does not lead to secondary structure formation^{15,19}. Furthermore, the same result also holds for the helices appearing in the collagen native state structure, which have a rather different geometry (in terms of local turn angles, residues per turn and pitch¹⁷) from average α -helices. In spite of these differences, we again obtained an average $f = 1.01 \pm 0.03$ (Fig. 3), very close to the optimal value. $\square$

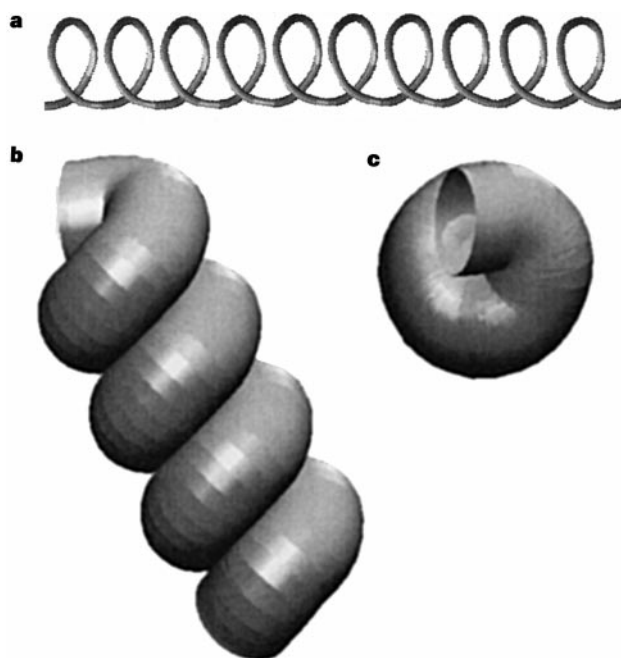


Figure 2 Optimal string with local constraint. The string has 65 points with a neighbour separation of unity. The local constraint was that each set of six consecutive beads had a radius of gyration less than 1. The value of f (see text for definition of the ratio) for this string is 0.9993. This result is quantitatively the same for a broad class of local

constraints. **a**, The 'bare' skeleton of the optimal helix connecting the discrete beads. **b, c**, Side and top views of the same helix inflated to its thickness. Note that there is no free space either between consecutive turns of the helix or in the plane perpendicular to the helix axis.

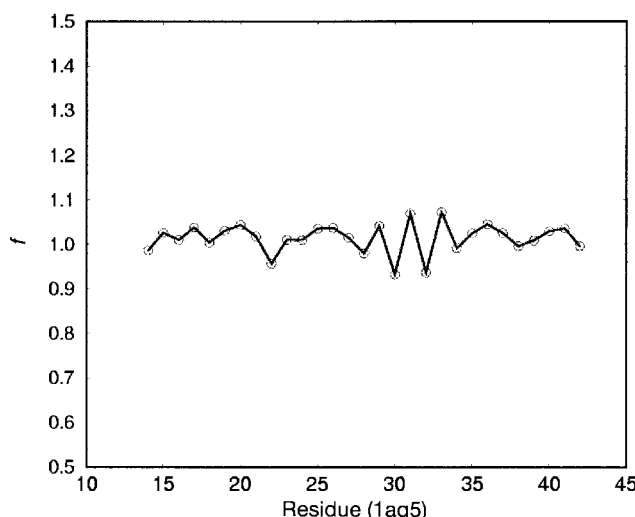


Figure 3 Packing of collagen helices. Plot of f values as a function of sequence position for a single collagen helix (only C_{α} coordinates were used to identify the protein backbone).

The sample plot for each of the three collagen chains would simply superimpose. We considered the residues 14–41 from the structure 1aq5 in the Protein Data Bank.

1. Sloane, N. J. A. Kepler's conjecture confirmed. *Nature* **395**, 435–436 (1998).
2. Mackenzie, D. Mathematics—proving the perfection of the honeycomb. *Science* **285**, 1339–1340 (1999).
3. Woodcock, L. V. Entropy difference between the face-centred cubic and the hexagonal close-packed crystal structures. *Nature* **385**, 141–143 (1997).
4. Car, R. Crystal structure—how hard spheres stack up. *Nature* **385**, 115–116 (1997).
5. Cipra, B. Mathematics—packing challenge mastered at last. *Science* **281**, 1267 (1998).
6. Stewart, I. Tight tins for round sardines. *Sci. Am.* Feb. 80–82 (1998).
7. Buck, G. & Orloff, J. A simple energy function for knots. *Topol. Appl.* **61**, 205–214 (1995).
8. Katritch, V. et al. Geometry and physics of knots. *Nature* **384**, 142–145 (1996).
9. Katritch, V., Olson, W. K., Pieranski, P., Dubochet, J. & Stasiak, A. Properties of ideal composite knots. *Nature* **388**, 148–151 (1997).
10. Gonzalez, O. & Maddocks, J. H. Global curvature, thickness and the ideal shapes of knots. *Proc. Natl Acad. Sci. USA* **96**, 4769–4773 (1999).
11. Buck, G. Four thirds power law for knots and links. *Nature* **392**, 238–239 (1998).
12. Cantarella, J., Kusner, R. B. & Sullivan, J. M. Tight knot values deviate from linear relations. *Nature* **392**, 237–238 (1998).
13. Rose, G. D. & Seltzer, J. P. A new algorithm for finding the peptide chain turns in a globular protein. *J. Mol. Biol.* **113**, 153–164 (1977).
14. Micheletti, C., Maritan, A. & Banavar, J. R. Protein structures and optimal folding from a geometrical variational principle. *Phys. Rev. Lett.* **82**, 3372–3375 (1999).
15. Maritan, A., Micheletti, C. & Banavar, J. R. Role of secondary motifs in fast folding polymers: a dynamical variational principle. *Phys. Rev. Lett.* **84**, 3009–3012 (2000).
16. Sali, A., Shakhnovich, E. & Karplus, M. How does a protein fold? *Nature* **369**, 248–251 (1994).
17. Creighton, T. E. *Proteins—Structures and Molecular Properties* 182–188 (Freeman, New York, 1993).
18. Yee, D. P., Chan, H. S., Havel, T. F. & Dill, K. A. Does compactness induce secondary structure in proteins? Study of poly-alanine chains computed by distance geometry. *J. Mol. Biol.* **241**, 557–573 (1994).
19. Succi, N. D., Bialek, W. S. & Onuchic, J. N. Properties and origins of protein secondary structure. *Phys. Rev.* **49**, 3440–3443 (1994).
20. Sokal, A. D. Monte Carlo methods for the self-avoiding walk. *Nucl. Phys.* **B47**, 172–179 (1996).

Acknowledgements

This work was supported by INFN, NASA and The Donors of the Petroleum Research Fund administered by the American Chemical Society. A.T. thanks the Physics Department of Università degli Studi di Padova, Padova, Italy, for its hospitality.

Correspondence and requests for materials should be addressed to J.R.B. (e-mail: jayanth@phys.psu.edu).

Evidence for a link between global lightning activity and upper tropospheric water vapour

Colin Price

Department of Geophysics and Planetary Sciences, Tel Aviv University, Levonon Road, Ramat Aviv 69978, Israel

Tropospheric water vapour is a key element of the Earth's climate, which has direct effects as a greenhouse gas, as well as indirect effects through interaction with clouds, aerosols and tropospheric chemistry. Small changes in upper-tropospheric water vapour have a much larger impact on the greenhouse effect than small changes in water vapour in the lower atmosphere¹, but whether this impact is a positive or negative feedback remains uncertain^{2–6}. The main challenge in addressing this question is the difficulty in monitoring upper-tropospheric water vapour globally over long timescales. Here I show that upper-tropospheric water-vapour variability and global lightning activity are closely linked, suggesting that upper-tropospheric water-vapour changes can be inferred from records of global lightning activity, readily obtained from observations at a single location on the Earth's surface. This correlation reflects the fact that continental deep-convective thunderstorms transport large amounts of water vapour into the upper troposphere and thereby dominate the variations of global upper-tropospheric water vapour while producing most of the lightning on Earth. As global lightning induces Schumann

resonances, an electromagnetic phenomenon in the atmosphere that can be observed easily at low cost, monitoring of these resonances might provide a convenient method for tracking upper-tropospheric water-vapour variability and hence contribute to a better understanding of the processes affecting climate change.

Both climate models and observations support the idea that higher temperatures will increase the amount of upper-tropospheric water vapour (UTWV)^{2–5}. Observations indicate that UTWV may already be increasing⁷. Although there are some claims to the contrary⁶, there is little evidence to support the argument for the existence of a negative water-vapour feedback. Climate models predict water-vapour increases in the upper troposphere of approximately 10% for every 1 K increase in temperature in that layer. Some climate models predict UTWV to increase by 20% for every 1 K increase in surface temperatures⁵. This sensitivity is greater than that predicted by the Clausius–Clapeyron equation (6% per 1 K at 300 K) since UTWV is influenced not only by temperature, but also by transport from the lower atmosphere. Furthermore, while tropical surface temperatures may increase by 2–3 °C in a warmer climate, the upper tropical troposphere is expected to warm by 6–7 °C (ref. 1). As a result, the water-vapour feedback could amplify the surface temperature change due to a doubling of carbon dioxide by 60% (ref. 8).

UTWV is transported aloft by deep convection, later to be redistributed zonally and meridionally in the upper atmosphere⁹. A significant fraction of the water mass that is cycled through these clouds is transported as liquid droplets and ice particles. The stronger the updrafts, the deeper these cloud and precipitation particles are transported into the upper levels of clouds. Not only does the intensity of the convection influence the volume of water transported aloft, but it simultaneously influences the electrification processes in these convective clouds¹⁰. Radar studies of deep tropical convective storms show significant differences in the vertical structure of storms with large lightning activity relative to those storms that are weakly electrified¹¹. In the case of storms with large lightning frequencies, the updrafts are strong enough to transport large amounts of supercooled droplets and graupel particles (soft hail) into the upper troposphere. The eventual sublimation of the large ice-filled anvil clouds results in a major source of water vapour into the upper troposphere^{12,13}. It is possible that the precipitation and sublimation of large ice particles from these anvils will also moisten the mid-troposphere¹⁴, while the advection of water vapour over the oceans¹¹ may result in the re-formation of new cirrus over the oceans¹⁵. Large amounts of subvisible cirrus occur near the tropical

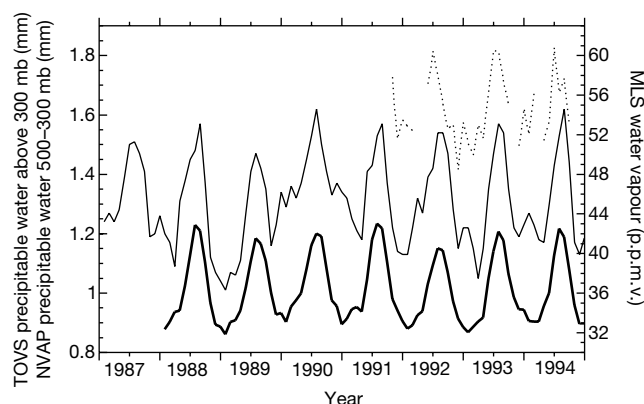


Figure 1 Seasonal variability of upper-tropospheric water vapour (UTWV). Monthly concentrations of water vapour at 215 mbar (MLS; dotted line), precipitable water above 300 mbar (TOVS; thin line) and precipitable water between 500–300 mbar (NVAP; thick line). The correlation between the NVAP and TOVS data sets (see text for description of data sets) is $r = 0.88$.

tropopause downwind from centres of tropical continental convection¹². The absorption of terrestrial radiation by these clouds results in heating and slow lifting, promoting the formation and persistence of the cirrus layer over the oceans.

It has already been demonstrated that enhanced tropical convection is associated with enhanced upper-tropospheric humidity¹⁶. Positive correlations are observed for a wide range of spatial and temporal scales. Furthermore, it has also been shown that the greenhouse effect is enhanced above storm tracks and deep convective storms owing to the transport of large amounts of water vapour into the upper troposphere^{16,17}. Simulations indicate that for a climate modelled with twice the present-day amount of CO₂, the largest percentage increase in water vapour occurs in the tropical upper troposphere, closely related to the location of the tropical deep convection¹. Here I present a new diagnostic for studying global UTWV variability. It involves using the global atmospheric electric circuit that is regulated by global lightning activity. The use of the global electric circuit to monitor climate variability has been suggested^{18–22}. However, these studies focused on using the global electric circuit and lightning activity to study tropical and global surface temperature changes.

I analysed UTWV from three different sources: the Upper Atmospheric Research Satellite (UARS) Microwave Limb Sounder (MLS)²³; the Television Infrared Observation Satellite (TIROS) Operational Vertical Sounder (TOVS) Pathfinder A (ref. 24); and the NASA Water Vapor Project (NVAP)²⁵. All these UTWV data sets show clear maxima over the tropical landmasses, exactly in the regions where deep convective clouds are highest²³, and where the lightning activity is maximum²⁶. On a seasonal timescale the continental tropical UTWV concentrations are 50–60% larger than over the adjacent oceanic regions in the intertropical convergence zone. Hence, the intensity of continental convection is important in determining the global lightning activity as well as the global UTWV concentrations. The opposite occurs for the lower-tropospheric water-vapour distributions which are dominated by evaporation over the oceans.

The monthly variability of these three UTWV data sets is shown in Fig. 1. The MLS data are from a limb sounder with measurements only at local sunrise and sunset, and represent the highest altitude observations (215 mbar). The TOVS data represent the precipitable water above 300 mbar, while the NVAP data represent the precipitable water within the layer 500–300 mbar. Although the data sets show different absolute values for UTWV, they all show a clear annual cycle with the Northern Hemisphere summer having larger concentrations of UTWV than the Southern Hemisphere summer. However, if we analyse one of these data sets in more detail (Fig. 2),

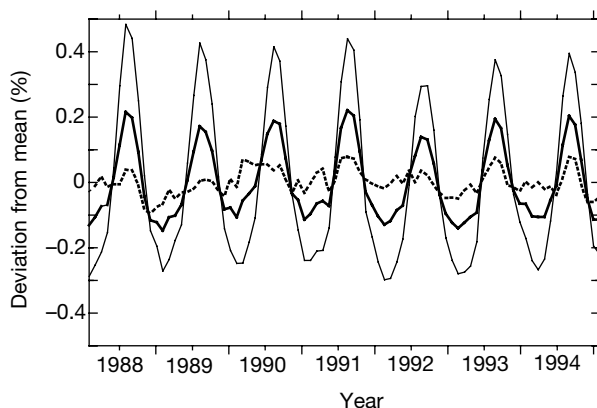


Figure 2 Variability of NVAP UTWV. The monthly deviations (per cent) from the mean for UTWV over land (thin line), oceans (dotted line) and the entire globe (thick line). The seven-year mean values are: land, ~0.4 mm; ocean, ~0.6 mm; globe, ~1 mm.

we notice that the majority of the annual variability is due to the continental UTWV, while the oceanic contribution is fairly constant throughout the year. The variability of the UTWV above land is approximately $\pm 35\%$, but over the oceans it is only $\pm 5\%$. Therefore, although the mean UTWV concentrations are always greater over the oceans (ocean mean, ~0.6 mm; land mean, ~0.45 mm), the variability of global UTWV is dominated by the convection over the continents. It is possible that some of the annual variability observed over the oceans is actually 'spillover' from continental deep-convective storms.

A relatively simple and cheap method for continuously observing global lightning variability is using the Schumann resonances²⁷. Each lightning discharge emits electromagnetic radiation at all frequencies and in all directions. The lower the frequency of the radiation, the less the attenuation of the electromagnetic waves in the atmosphere, and the greater the propagation distance. At extremely low frequencies (ELF, $1 \text{ Hz} < f < 100 \text{ Hz}$) the radiation can propagate a few times around the globe before dissipating. This is achieved by electromagnetic waves being trapped in the earth-ionosphere waveguide. The resonant frequencies are excited by lightning discharges in the lower atmosphere, and as there are approximately 50–100 lightning flashes per second around the globe, the variability of the intensity of the Schumann resonances represents a continuous measure of the variability of global lightning activity²⁸. Furthermore, simultaneous measurements on opposite sides of the globe show remarkable agreement (Fig. 3), implying that a single station can be used to track global lightning variability.

One station that has been monitoring 10 Hz ELF radiation since 1985 is located at Arrival Heights, Antarctica²⁹. This station records the horizontal magnetic field components of the ELF radiation produced by lightning. The data are sampled 600 times per minute, and later averaged into daily and monthly means. Comparing the monthly mean ELF observations from Antarctica with the globally integrated monthly mean NVAP precipitable water values between 500–300 mbar (Fig. 4) it is clear that the continental convection is driving both the global UTWV fluctuations as well as the global lightning activity ($r = 0.85$). If we use only the continental values of UTWV, the correlation increases slightly to $r = 0.88$. (The correlations between the ELF and the global TOVS and MLS UTWV data sets are 0.65 and 0.68 respectively). The agreement between the ELF data and the NVAP data is similar to that between the water-vapour data sets themselves (Fig. 2). The sensitivity of this relationship implies that for every 0.25 mm increase in UTWV (1% of total column), the ELF signal will increase by approximately 25%. This could result from either an increase in the amount of deep convection

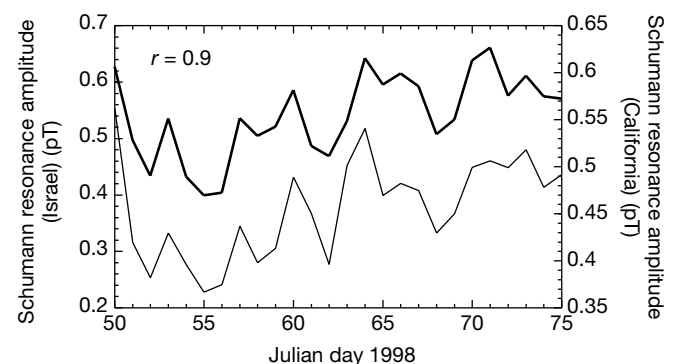


Figure 3 Schumann resonance as a global parameter. The amplitude of the 8-Hz horizontal magnetic field observed simultaneously at a station in Israel (thick line) and a station in California (thin line) over a 25-day period in 1998. Although the observations at the two sites were totally independent, the daily correlation between them is 0.9. (1 pT = 10^{-12} T.)

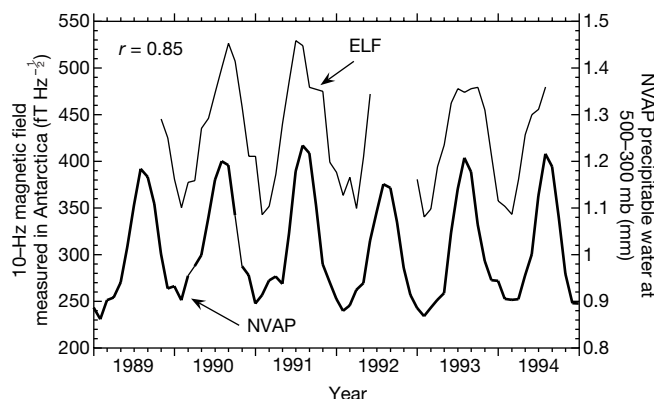


Figure 4 Monthly extremely low frequency (ELF)–UTWV correlations. The monthly mean variability of the 10 Hz magnetic field intensity measured in Antarctica (thin line) compared with the monthly variability of the NVAP UTWV (thick line). (1fT = 10^{-15} T.)

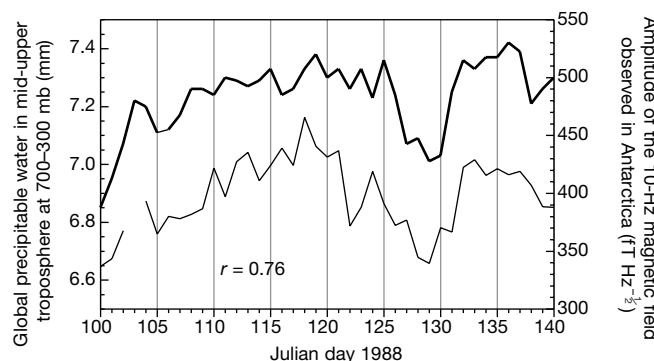


Figure 5 Daily ELF–UTWV correlations. The daily mean variability of the global NVAP precipitable water above 700 mbar from 9 April to 19 May 1998 (thick line), together with the 10-Hz magnetic field amplitude measured in Antarctica (thin line). (1fT = 10^{-15} T.)

or an increase in the intensity of deep convection.

Although the daily UTWV data are rather noisy, there are periods when good agreement is found between the ELF data and the NVAP UTWV (Fig. 5). I note that the ELF maxima and minima often precede the water-vapour maxima and minima on a daily basis. It is possible that the thunderstorms that produce the lightning activity result in UTWV maxima one or two days after the lightning activity peaks.

In addition to the sampling problems related to the various UTWV data sets, the ELF data from Antarctica are available only for the frequency of 10 Hz during the above periods. Ideally, the 8-Hz signal (Fig. 3) would be a better indicator of global lightning activity. Nevertheless, the magnetic field data at 10 Hz also represent global lightning activity, and the daily and monthly temporal variability is almost the same. Hence, it is believed that the relationships presented here will only improve when better data sets become available.

The agreement between the variability of ELF intensities and global UTWV concentrations suggests that single-station measurements of the Schumann resonances could supply a cheap, continuous, long-term measure of the variability of UTWV. Although this study deals primarily with the annual variability of UTWV, climate models indicate that future changes in deep convection and lightning activity may be larger over the continents than over the oceans³⁰, as a result of greenhouse warming. Therefore, the Schumann resonances may also provide a useful tool to monitor future changes in upper-tropospheric water-vapour as a result of global climate change. □

Received 6 April; accepted 19 May 2000.

1. Hansen, J. *et al.* in *Climate Processes and Climate Sensitivity* (eds Hansen, J. E. & Takahashi, T.) 130–162 (American Geophysical Union, Washington DC, 1984).

2. Rind, D. *et al.* Positive water vapour feedback in climate models confirmed by satellite data. *Nature* **349**, 500–502 (1991).
3. Del Genio, A. D., Kovari, W. Jr. & Yao, M. S. Climatic implications of the seasonal variations of upper troposphere water vapor. *Geophys. Res. Lett.* **21**, 2701–2704 (1994).
4. Sun, D. Z. & Held, I. M. A comparison of modeled and observed relationships between interannual variations of water vapor and temperature. *J. Clim.* **9**, 665–675 (1996).
5. Rind, D. Just add water vapor. *Science* **281**, 1152–1153 (1998).
6. Lindzen, R. S. Some coolness concerning global warming. *Bull. Am. Meteorol. Soc.* **71**, 288–299 (1990).
7. Oltmans, S. J. & Hofmann, D. J. Increase in lower-stratospheric water vapor at a mid-latitude Northern Hemisphere site from 1981–1994. *Nature* **374**, 146–148 (1995).
8. Houghton, J. J. *et al.* (eds) *Climate Change 1995: IPCC Assessment Report* (Cambridge Univ. Press, Cambridge, 1996).
9. Newell, R. E. *et al.* Walker circulation and tropical upper tropospheric water vapor. *J. Geophys. Res.* **101**, 1961–1974 (1996).
10. Takahashi, T. Near absence of lightning in torrential rainfall producing Micronesian thunderstorms. *Geophys. Res. Lett.* **17**, 2381–2384 (1990).
11. Williams, E. R. *et al.* A radar and electrical study of tropical “hot towers”. *J. Atmos. Sci.* **49**, 1386–1395 (1992).
12. Kent, G. S. *et al.* Surface temperature related variations in tropical cirrus cloud as measured by SAGE II. *J. Clim.* **8**, 2577–2594 (1995).
13. Houze, R. A. Observed structure of mesoscale convective systems and implications for large-scale heating. *Q. J. R. Meteorol. Soc.* **115**, 425–430 (1989).
14. Sun, D.-Z. & Lindzen, R. S. Distribution of tropical water vapor. *J. Atmos. Sci.* **50**, 1643–1660 (1993).
15. Jensen, E. J. *et al.* High humidities and subvisible cirrus near the tropical tropopause. *Geophys. Res. Lett.* **26**, 2347–2350 (1999).
16. Sodan, B. J. & Fu, R. A satellite analysis of deep convection, upper-tropospheric humidity, and the greenhouse effect. *J. Clim.* **8**, 2333–2339 (1995).
17. Hu, H. & Liu, W. T. The impact of upper tropospheric humidity from Microwave Limb Sounder on the midlatitude greenhouse effect. *Geophys. Res. Lett.* **25**, 3151–3154 (1998).
18. Williams, E. R. The Schumann Resonance: a global tropical thermometer. *Science* **256**, 1184–1186 (1992).
19. Price, C. Global surface temperatures and the atmospheric electric circuit. *Geophys. Res. Lett.* **20**, 1363–1366 (1993).
20. Williams, E. R. Global circuit response to seasonal variations in global surface air temperature. *Mon. Weath. Rev.* **122**, 1917–1929 (1994).
21. Reeve, N. & Tuomi, R. Lightning activity as an indicator of climate change. *Q. J. R. Meteorol. Soc.* **125**, 893–903 (1999).
22. Markson, R. & Price, C. Ionospheric potential as a proxy index for global temperatures. *Atmos. Res.* **51**, 309–314 (1999).

23. Read, W. G. *et al.* Upper-tropospheric water vapor from UARS MLS. *Bull. Am. Met. Soc.* **76**, 2381–2389 (1995); also at (http://mls.jpl.nasa.gov/joe/h2o_uptrop_dif_jja.html).
24. Susskind, J. *et al.* Characteristics of the TOVS Pathfinder Path A dataset. *Bull. Am. Met. Soc.* **78**, 1449–1472 (1997).
25. Randel, D. L. *et al.* A new global water vapor dataset. *Bull. Am. Meteorol. Soc.* **77**, 1233–1246 (1996).
26. Latham, J. & Christian, H. Satellite measurements of global lightning. *Q. J. R. Meteorol. Soc.* **124**, 1771–1773 (1998).
27. Schumann, W. O. Über die strahlungslosen Eigenschwingungen einer leitenden Kugel, die von einer Luftschicht und einer Ionosphärenhülle umgeben ist. *Z. Naturforsch.* **7a**, 149–162 (1952).
28. Heckman, S. J., Williams, E. & Boldi, B. Total global lightning inferred from Schumann Resonance measurements. *J. Geophys. Res.* **103**, 31775–31779 (1998).
29. Fraser-Smith, A. C. *et al.* in *Environmental and Space Electrodynamics* (ed. Kikuchi, H.) 191–200 (Springer, Tokyo, 1991).
30. Price, C. *et al.* Possible implications of global climate change on global lightning distributions and frequencies. *J. Geophys. Res.* **99**, 10823–10831 (1994).

Acknowledgements

I would like to thank T. Fraser-Smith for the use of the Antarctica ELF data used in this study and M. Füllekrug for making available the California data in Fig. 3 on the Internet. Part of this research was supported by the Israel-US Binational Science Foundation, and the European Community International Association for the promotion of co-operation with scientists from the New Independent States of the former Soviet Union (INTAS) programme.

Correspondence and requests for materials should be addressed to C. P. (e-mail: cpriace@flash.tau.ac.il).

Melting of the Earth's lithospheric mantle inferred from protactinium–thorium–uranium isotopic data

Yemane Asmerom*, Hai Cheng†, Rebecca Thomas†, Marc Hirschmann† & R. Lawrence Edwards†

* Department of Earth and Planetary Sciences, University of New Mexico, Albuquerque, New Mexico 87131, USA

† Department of Geology and Geophysics, University of Minnesota, Minneapolis, Minnesota 55455, USA

The processes responsible for the generation of partial melt in the Earth's lithospheric mantle and the movement of this melt to the Earth's surface remain enigmatic, owing to the perceived difficulties in generating large-degree partial melts at depth and in transporting small-degree melts through a static lithosphere¹. Here we present a method of placing constraints on melting in the lithospheric mantle using ²³¹Pa–²³⁵U data obtained from continental basalts in the southwestern United States and Mexico. Combined with ²³⁰Th–²³⁸U data^{2,3}, the ²³¹Pa–²³⁵U data allow us to constrain the source mineralogy and thus the depth of melting of these basalts. Our analysis indicates that it is possible to transport small melt fractions—of the order of 0.1%—through the lithosphere, as might result from the coalescence of melt by compaction⁴ owing to melting-induced deformation⁵. The large observed ²³¹Pa excesses require that the timescale of melt generation and transport within the lithosphere is small compared to the half-life of ²³¹Pa (~32.7 kyr). The ²³¹Pa–²³⁰Th data also constrain the thorium and uranium distribution coefficients for clinopyroxene in the source regions of these basalts to be within 2% of one another, indicating that in this setting ²³⁰Th excesses are not expected during melting at depths shallower than 85 km.

Basalts from old continental settings are ideally suited for the study of lithospheric melting processes because they allow discrimination between lithospheric and asthenospheric mantle melts using Nd isotopes². The samples for this study are well-characterized North

American continental basalts for which Nd and Th isotopic compositions and trace-element abundances have previously been reported^{2,3}. The two groups of samples consist of: (1) asthenospheric mantle-derived alkali basalts (high ϵ_{Nd} values of around +5) from the Pinacate volcanic field, Mexico, and the Potrillo volcanic field in the southern Rio Grande Rift, USA; and (2) lithospheric mantle-derived (low ϵ_{Nd} values of around 0) alkali basalts from the San Francisco volcanic field, Colorado Plateau, and a tholeiite from the Zuni Bandera volcanic field in the Colorado Plateau–Rio Grande transition (Fig. 1).

As shown previously^{2,3}, the ²³⁰Th/²³⁸U enrichments in continental basalts correlate with their inferred mantle sources. The lavas with low ϵ_{Nd} values, thought to be derived from the lithospheric mantle, have equilibrium (²³⁰Th/²³⁸U) values, whereas lavas with high ϵ_{Nd} values, presumed to be derived from the asthenosphere, have large ²³⁰Th enrichments. Initial (²³¹Pa/²³⁵U) values in asthenospheric alkalic lavas (Table 1) fall in a narrow range from 2.33 ± 14 to 2.44 ± 15 (the uncertainty is mostly related to uncertainty in age), whereas the lithospheric mantle values are 2.00 ± 3 and 2.02 ± 2 for the alkali basalts and 1.44 ± 3 for the tholeiite.

The salient feature of the Pa data is that despite significant differences in geodynamic setting and other geochemical characteristics, all of the samples, and particularly the alkali olivine basalts, have significant (²³¹Pa/²³⁵U) excesses. This is true both for samples that show large (²³⁰Th/²³⁸U) excesses and for those that are in secular equilibrium with respect to (²³⁰Th/²³⁸U). A continental basalt from the Snake River Plain that has substantial Pa excess [²³¹Pa/²³⁵U = 1.7] but that is near secular equilibrium with respect to ²³⁰Th—(²³⁰Th/²³⁸U) = 1.04—was also reported⁶.

Alkali olivine basalts (AOBs) have significantly greater Pa enrichments than the one tholeiite analysed in this study. Similar differences between tholeiite and alkali basalts are observed in other settings⁶. As the AOBs are probably the result of smaller degrees of melting than the tholeiite^{2,3}, this is the expected relationship if Pa enrichment is caused by melting. The limited available data do not completely exclude exotic enrichment processes such as metasomatism, but here we will assume that the data are most probably explained by melting processes.

The coupled ²³⁰Th and ²³¹Pa excesses observed in the alkali basalts from the Pinacate volcanic field and the Potrillo volcanic field, which are derived from asthenospheric mantle^{2,3,7–9}, may be explained by the melting of upwelling depleted (high ϵ_{Nd}) mantle, similar to the reason for excesses in MORB and ocean island

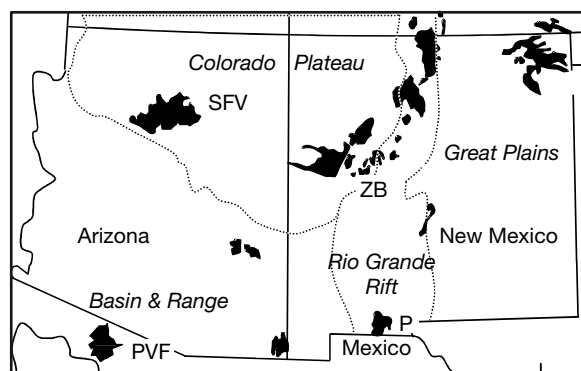


Figure 1 A general tectonic map of the study region. Shown is the Colorado Plateau, a stable part of cratonic North America (with a thick lithosphere), bounded by the Rio Grande Rift and the Basin and Range Province, USA and Mexico (regions with attenuated lithosphere). The four sample localities are: (1) the Pinacate volcanic field (PVF), Mexico (samples 92-03 and 92-04); (2) the Potrillo volcanic field (P), southern Rio Grande Rift, USA (95-06), both with highly attenuated lithosphere; (3) the San Francisco volcanic field (SFV), Colorado Plateau (92-07, 92-08); and (4) the Zuni Bandera volcanic field (ZB, 95-04), a transitional zone with attenuated lithosphere. Other Cenozoic volcanic fields are shown for context.

basalts^{10–12}. The large ^{231}Pa enrichment ($(^{231}\text{Pa}/^{235}\text{U})$ values of up to 2.00) with no ^{230}Th enrichment in the lithosphere-derived basalts is a surprising discovery. As we will show below, it is not possible to generate significant ^{231}Pa enrichments while still maintaining secular equilibrium ^{230}Th values in a garnet peridotite source using either batch melting or dynamic melting models. Thus, these data require melting in a source without garnet. This conclusion, based on combined ^{231}Pa and ^{230}Th data, is supported by previous inferences³, attributing the lack of ^{230}Th enrichment to melting within the spinel lherzolite field. Moreover, xenolith and geophysical data from this part of the Colorado Plateau are supportive of a spinel lherzolite lithospheric mantle. In the Colorado Plateau the depth to the

lithosphere–asthenosphere boundary varies from less than 60 km near the margins of the Plateau to 110 km at the centre¹³. Although xenoliths from the central part of the Plateau include garnet peridotites, those closer to the margin in settings similar to the San Francisco volcanic field consist of spinel lherzolites¹⁴. The lithospheric mantle at the study site is clearly within the stability field of spinel lherzolite. Melting is likely to occur at the base of the lithosphere, as a result of basal heating. As a result, the portion of the lithosphere that had partially melted is probably located beneath the present-day seismic lithosphere–asthenosphere boundary¹³ (about 60 km), but above the stability of garnet peridotite¹⁵ (less than 85 km).

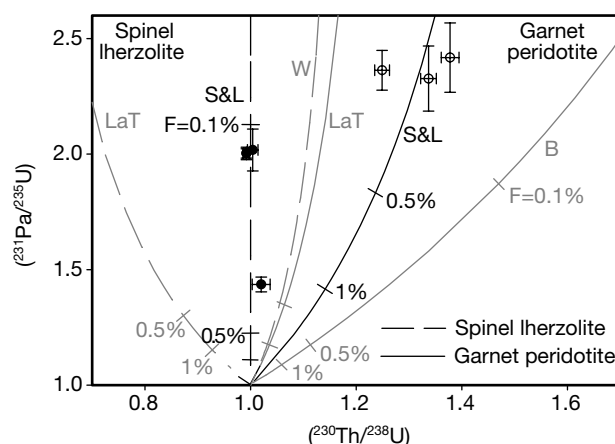


Figure 2 $(^{231}\text{Pa}/^{235}\text{U})$ versus $(^{230}\text{Th}/^{238}\text{U})$ diagram, showing our age-corrected values. Lithospheric mantle-derived basalts, with Nd isotopic values near bulk Earth, $\epsilon_{\text{Nd}} \approx 0$, are shown in filled symbols^{2,3} while those with high ϵ_{Nd} values, derived from an asthenospheric source^{2,3,7–9}, are shown with unfilled symbols. The plotted activity ratios are decay-corrected (initial) values. The uncertainties reflect both measurement errors and uncertainties associated with ages of the samples. F , degree of melting. All samples are alkali basalts, except for one tholeiitic basalt (95-04) with $(^{231}\text{Pa}/^{235}\text{U}) = 1.4$. We note that all of the samples, including the tholeiite, are enriched in ^{231}Pa . Samples with low ϵ_{Nd} values are in secular equilibrium with respect to ^{230}Th , whereas those with high ϵ_{Nd} values have large ^{230}Th enrichments. The difference in ^{230}Th enrichment is attributed to difference in source mineralogy (spinel peridotite versus garnet peridotite, respectively^{2,3}). Also shown are ^{231}Pa – ^{230}Th enrichment trends for simple batch melting of spinel lherzolite (dashed lines) and garnet peridotite (solid lines) sources, using some of the

published U–Th D -values for clinopyroxene (cpx) and garnet (gt). There are no experimental data for the D -value for Pa; it is inferred to be in the range of 10^{-5} (see ref. 18). S&L, Salters and Longhi¹⁹ (dark lines); W, Wood *et al.*²⁰; B, Beattie¹⁶; LaT, LaTourette and Burnett²¹ and LaTourette *et al.*¹⁷ D -value determinations by other workers fall in between the extreme values. Considering these models, the only way to get large ^{231}Pa enrichments while maintaining secular equilibrium ^{230}Th is to melt a spinel lherzolite source with $D_{\text{Pa}}^{\text{cpx}}/D_{\text{Pa}}^{\text{gt}} \approx 1$, consistent with the data of ref. 19. We used $D_{\text{Pa}}^{\text{cpx}}$ and $D_{\text{Pa}}^{\text{gt}}$ for clinopyroxene with a wollastonite mole fraction of about 0.5. The bulk composition of clinopyroxene is estimated from analyses of lithospheric mantle-derived clinopyroxene from the study region²⁷. The alkali basalts with no ^{230}Th enrichment require melting of a spinel lherzolite with F values in the range of 0.1%, and larger F values (about 0.3%) for tholeiite. The samples with coupled ^{231}Pa – ^{230}Th enrichments fit the simple batch melting of a garnet peridotite ($F \approx 0.4\%$), although dynamic melting (Fig. 3) is most likely to be relevant here.

Table 1 Protactinium, thorium and uranium data

Sample	Age (kyr)*	^{238}U (ng g ⁻¹)	$(^{230}\text{Th}/^{238}\text{U})$	$(^{230}\text{Th}/^{238}\text{U})_{\text{t}}$ †	Pa (fg g ⁻¹)	$(^{231}\text{Pa}/^{235}\text{U})$	$(^{231}\text{Pa}/^{235}\text{U})_{\text{t}}$ ‡	$(^{234}\text{U}/^{238}\text{U})$
Asthenospheric mantle melts ($\epsilon_{\text{Nd}} \approx 5$)								
95-06 (II) [P]	20	1,127 ± 1	1.207 ± 13	1.249 ± 14	699 ± 24	1.893 ± 64	2.36 ± 86	1.006 ± 6
95-06 (I)		1,126 ± 1			672 ± 14	1.823 ± 38		0.997 ± 5
92-03 (II) [PVF]	~10	904 ± 1	1.307 ± 7	1.337 ± 15	615 ± 11	2.074 ± 36	2.327 ± 141	1.001 ± 6
92-03 (I)		905 ± 1			626 ± 44	2.110 ± 148		1.006 ± 5
92-04 (II) [PVF]	~10	848 ± 1	1.345 ± 9	1.378 ± 17	597 ± 9	2.148 ± 33	2.418 ± 150	1.003 ± 9
92-04 (I)		848 ± 1			596 ± 39	2.146 ± 141		1.007 ± 7
Lithospheric mantle melts ($\epsilon_{\text{Nd}} \approx 0$)								
92-07 [SFV]	0.80	1,095 ± 1	0.992 ± 5	0.992 ± 5	712 ± 10	1.986 ± 27	2.003 ± 27	0.999 ± 7
92-07§		1,102			714		2.003 ± 20	
92-07§		1,102			714		1.99 ± 20	
92-08 (II) [SFV]	0.80	1,030 ± 1	1.004 ± 10	1.004 ± 10	676 ± 31	2.001 ± 91	2.001 ± 91	1.002 ± 9
92-08 (I)		1,028 ± 1			668 ± 62	1.983 ± 185		
95-04 [ZB]	3.0	697 ± 1	1.019 ± 16	1.020 ± 17	276 ± 6	1.409 ± 32	1.436 ± 32	0.997 ± 8
Table Mountain Latite Standard								
Minnesota data		10,660 ± 12	1.002 ± 5		3,500 ± 25	1.002 ± 7		1.002 ± 1
LANL data†		10,425	1.003 ± 7		3,415	≡ 1.000		
		10,474	0.996 ± 8		3,433	≡ 1.000		

Th isotopic analyses were done at the University of New Mexico² and the University of Minnesota³. The Pa chemistry and mass spectrometry were done at the University of Minnesota. See refs 28 and 29 for Pa separation and mass spectrometry, and refs 2 and 3 for silicate dissolution. Ratios in parentheses are activity ratios. Uncertainties on measured ratios are 2σ of the mean analytical errors. Approximate ϵ_{Nd} values are from refs 2 and 3, in which measurements for all samples except 92-03 can be found.

* See refs 2 and 3 for sources of ages.

† Uncertainties in the initial ratios reflect both analytical errors of the measured values, and age uncertainties. The large errors on the initials for 92-03 and 92-04 are due to assigning 50% error on the ages. Given the large Pa excesses in the samples, any uncertainties in initial ratios are not crucial to the interpretations in this paper. The markedly lower $(^{231}\text{Pa}/^{235}\text{U})$ value for the tholeiitic basalt (95-04) is not related to post-eruption decay because the young age of the tholeiite, 3,000 years, is well constrained from ^{14}C and ^3He dating.

‡ The $(^{231}\text{Pa}/^{235}\text{U})$ ratio data for the Los Alamos Table Mountain Latite²⁸ is fixed at 1.000 as it was used to calibrate their spike.

§ U and Pa in 92-07 were analysed at Los Alamos National Labs (LANL) in replicate⁶. See refs 3 and 29 for decay constants of U, Th and Pa isotopes. See legend to Fig. 1 for abbreviations in square brackets.

The observed Pa and Th isotope results can be compared to simple models for melting of peridotite. In Fig. 2 the data are compared to the results of batch melting of garnet peridotite (12% clinopyroxene (cpx), 8% garnet) and spinel lherzolite (15% cpx) lithologies, using a variety of published experimental Th and U D -values and estimates of Pa D -values^{16–19}. (The D value (distribution coefficient) is the ratio of the concentration of an element in a solid phase over its concentration in melt.) The calculations show that the lavas with large ^{231}Pa excesses but no ^{230}Th enrichment are inconsistent with partial melting of garnet peridotite. Using any of the available D -values, when melting a garnet peridotite, any significant enrichment in ^{231}Pa results in a coupled enrichment of ^{230}Th . Figure 2 also shows that even if the source is free of garnet (spinel lherzolite) only the D -values of ref. 19 give melts with large ^{231}Pa enrichment and no Th–U fractionation, within measurement error; ($^{230}\text{Th}/^{238}\text{U}$) $\approx$ 1.

The ratio of experimentally determined U–Th D -values for clinopyroxene, $D_{\text{Th}}^{\text{cpx}}/D_{\text{U}}^{\text{cpx}}$, vary from less than one²⁰ to as high as 2.2 (ref. 21), and vary with composition¹⁸ and pressure²⁰. Our combined Th–Pa data may provide a robust constraint on $D_{\text{Th}}^{\text{cpx}}/D_{\text{U}}^{\text{cpx}}$. For simple batch melting of a source initially in secular equilibrium, assuming that $D_{\text{Pa}} \approx 0$ and degree of melting, F , is small (about 0.1%), $\frac{D_{\text{Th}}}{D_{\text{U}}} = \frac{(B-A)}{A(B-1)}$, where $A = (\frac{^{230}\text{Th}}{^{238}\text{U}})$ and $B = (\frac{^{231}\text{Pa}}{^{235}\text{U}})$. For alkali basalts derived from the lithosphere, accounting for analytical errors on the two samples (92-07, 92-08) analysed, the weighted average of the alkali basalt source region $D_{\text{Th}}/D_{\text{U}}$ implied by the Pa–Th–U data is 1.012 ± 0.012 (2σ). The Pa–Th–U data for the one tholeiite measured indicates a source-region $D_{\text{Th}}/D_{\text{U}}$ value of 0.935 ± 0.106 . The combined data yield a $D_{\text{Th}}/D_{\text{U}}$ value of 1.011 ± 0.012 (assuming that the tholeiite and alkali basalts share the same source region).

We observed lavas ranging in composition from tholeiites to alkali basalts that show no measurable Th excesses, but large ^{231}Pa enrichment, suggesting that the $D_{\text{Th}}^{\text{cpx}}/D_{\text{U}}^{\text{cpx}}$ is within 1% of unity, at least for these settings. Suggestions, based on theoretical considerations and limited experimental data²⁰, that $D_{\text{Th}}^{\text{cpx}}/D_{\text{U}}^{\text{cpx}}$ may be lower than one at pressures similar to those in the source regions of the

samples we derived from the lithospheric mantle, are not supported by our data.

For the lithospheric mantle-derived melts (no ^{230}Th enrichment) very small degrees of melting (F -values in the range of 0.1%) are required to generate the observed ^{231}Pa excesses (Fig. 2, curve from ref. 19). Such small extents of melting are broadly compatible with extents of melting for these lavas implied by concentrations of other incompatible trace elements such as La and Nd in the lavas [$\text{La} \approx 100$ and $\text{Nd} \approx 30$ times the chondritic values]². For example, using the D -values of ref. 22, these extents of melting imply enrichments relative to their source of about 100–120 for La and about 35 for Nd. Although more complicated processes than batch melting^{23,24} may be responsible for the observed ^{231}Pa excesses, all viable models that explain these data would require melting or melt transport at porosities of around 0.1% (Fig. 3).

One possibility is that significant separation between Pa and U may be enhanced by chromatographic separation during percolation through the lithosphere²⁵. In such a process the melts ascend but the solid mantle does not upwell, and thus percolating melts are able to exchange with mantle that has not melted recently and therefore may have a larger Pa/U ratio than that of upwelling, partially melting mantle. If it occurs at sufficiently low porosity, such percolative exchange may be efficient at generating ^{231}Pa over ^{235}U excesses.

For samples derived from the asthenospheric mantle with coupled Th and Pa enrichment, both static and dynamic melting models^{23,24,26} may be applicable. Batch melting of garnet peridotite with a maximum extent of melting of about 0.4% can approximate the calculated Pa–Th–U systematics of these rocks (Fig. 2). More importantly, partial melting of upwelling mantle is most probably approximated by ingrowth models, allowing for larger (about 3%) effective degrees of partial melting (Fig. 3). For example, the data are well matched by some dynamic models^{23,26}, and by an equilibrium porous flow model²⁴ as illustrated in Fig. 3, provided that the upwelling rate is on the order of 1 cm yr^{-1} or less.

Our data place strong constraints on the processes by which melting and melt extraction occur within the heated lithospheric mantle. The large observed ^{231}Pa excesses require that the timescale of melt generation and transport within the lithosphere is small compared to the half-life of ^{231}Pa (about 32.7 kyr). The data strongly suggest that it is possible to mobilize small degrees of partial melt, in the range of 0.1%. Such small melt fractions may coalesce and move by compaction⁴ if the lithosphere experiences local deformation such as shear associated with melting, as is believed to occur in the asthenosphere⁵. □

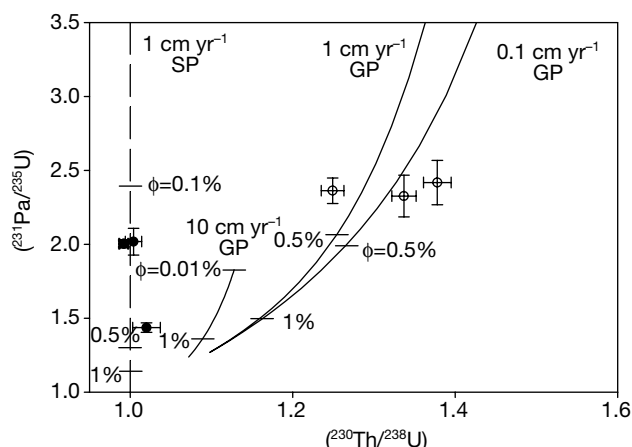


Figure 3 Results from dynamic melting calculations using an equilibrium porous flow model²⁴ and D values from ref. 19. We have used a 10-km melting column at a maximum F value of 3%. Porosities (ϕ) are shown as tick marks. The model of ref. 24 gives larger enrichments for a given porosity values at upwelling rates of more than 0.1 cm yr^{-1} than the models of refs 23 and 26 (not shown); these two models are indistinguishable at upwelling rates of less than 0.1 cm yr^{-1} . Dynamic melting of a garnet peridotite source is probably appropriate for the asthenosphere derived lavas, similar to MORB, in which melting occurs as a result of decompression melting. Dynamic melting results in larger enrichment compared to static melting, such as batch melting, for equivalent F values. The initial ratios are plotted to show that the degrees of enrichment seen in the samples are possible using physically reasonable upwelling rates and melt porosities. SP, spinel lherzolite; GP, garnet peridotite.

Received 26 October 1999; accepted 8 June 2000.

1. Leeman, W. P. & Harry, D. L. A binary source model for extension-related magmatism in the Great Basin, western North America. *Science* **262**, 1550–1554 (1993).
2. Asmerom, Y. The Th–U fractionation and mantle structure. *Earth Planet. Sci. Lett.* **166**, 163–175 (1999).
3. Asmerom, Y. & Edwards, R. L. U-series isotope evidence for the origin of continental basalts. *Earth Planet. Sci. Lett.* **134**, 1–7 (1995).
4. Kohlstedt, D. L. & Zimmerman, M. E. Rheology of partially molten mantle rocks. *Annu. Rev. Earth Planet. Sci.* **24**, 41–62 (1996).
5. McKenzie, D. The generation and compaction of partially molten rock. *J. Petrol.* **25**, 713–765 (1984).
6. Pickett, D. A. & Murrell, M. T. Observations of $^{231}\text{Pa}/^{235}\text{U}$ disequilibrium in volcanic rocks. *Earth Planet. Sci. Lett.* **148**, 259–271 (1997).
7. Perry, E., Baldrige, W. & DePaolo, D. Role of asthenosphere and lithosphere in the genesis of Late Cenozoic basaltic rocks from the Rio Grande rift and adjacent regions of the southwestern United States. *J. Geophys. Res.* **92**, 9193–9213 (1987).
8. Menzies, M., Kyle, P., Jones, M. & Ingram, G. Enriched and depleted source components for tholeiitic and alkaline lavas from Zuni-Bandera, New Mexico: inferences about intraplate processes and stratified lithosphere. *J. Geophys. Res.* **96**, 13645–13671 (1991).
9. Lynch, D. J., Musselman, T. E., Gutmann, J. T. & Patchett, P. J. Isotopic evidence for the origin of Cenozoic volcanic rocks in the Pinacate volcanic field, northwestern Mexico. *Lithos* **29**, 295–302 (1993).
10. Goldstein, S. J., Murrell, M. T. & Williams, R. W. ^{231}Pa and ^{230}Th chronology of mid-ocean ridge basalts. *Earth Planet. Sci. Lett.* **115**, 151–159 (1993).
11. Lundstrom, C. C., Gill, J., Williams, Q. & Hanan, B. B. Investigating solid mantle upwelling beneath mid-ocean ridges using U-series disequilibrium. II. A local study at 33° S Mid-Atlantic Ridge. *Earth Planet. Sci. Lett.* **157**, 167–181 (1998).

12. Bourdon, B., Joron, J.-L., Christelle, C.-I. & Allegre, C. J. U-Th-Pa-Ra systematics for the Grande Comore volcanics; melting processes in an upwelling plume. *Earth Planet. Sci. Lett.* **164**, 119–133 (1998).
13. Thompson, G. A. & Zoback, M. L. Regional geophysics of the Colorado Plateau. *Tectonophysics* **61**, 149–181 (1979).
14. Alibert, C. Peridotite xenoliths from western Grand Canyon and the Thumb; a probe into the subcontinental mantle of the Colorado Plateau. *J. Geophys. Res.* **99**, 21605–21620 (1994).
15. Robinson, J. A. & Wood, B. J. The depth of the spinel to garnet transition at the peridotite solidus. *Earth Planet. Sci. Lett.* **164**, 277–284 (1998).
16. Beattie, P. Uranium–thorium disequilibria and partitioning on melting of garnet peridotite. *Nature* **363**, 63–65 (1993).
17. LaTourrette, T. Z., Kennedy, A. K. & Wasserburg, G. J. Thorium–uranium fractionation by garnet: Evidence for a deep source and rapid rise of oceanic basalts. *Science* **261**, 739–742 (1993).
18. Lundstrom, C. *et al.* Compositional controls on the partitioning of U, Th, Ba, Pb, Sr and Zr between clinopyroxene and haplobasaltic melts: implications for uranium series disequilibria in basalts. *Earth Planet. Sci. Lett.* **128**, 407–423 (1994).
19. Salters, V. J. M. & Longhi, J. Trace element partitioning during the initial stages of melting beneath mid-ocean ridges. *Earth Planet. Sci. Lett.* **166**, 15–30 (1999).
20. Wood, B. J., Blundy, J. D. & Robinson, J. A. The role of clinopyroxene in generating U-series disequilibrium during mantle melting. *Geochim. Cosmochim. Acta* **63**, 1613–1620 (1999).
21. LaTourrette, T. Z. & Burnett, D. S. Experimental determination of U and Th partitioning between clinopyroxene and natural and synthetic basaltic liquid. *Earth Planet. Sci. Lett.* **110**, 227–244 (1992).
22. Hirschmann, M. M. & Stolper, E. M. A possible role for garnet pyroxenite in the origin of the “garnet signature” in MORB. *Contrib. Mineral. Petrol.* **124**, 185–208 (1986).
23. Williams, R. W. & Gill, J. B. Effect of partial melting on uranium decay series. *Geochim. Cosmochim. Acta* **53**, 1607–1619 (1989).
24. Spiegelman, M. & Elliott, T. Consequences of melt transport for uranium series disequilibrium in young lavas. *Earth Planet. Sci. Lett.* **118**, 1–20 (1993).
25. Navon, O. & Stolper, E. Geochemical consequences of melt percolation: the upper mantle as a chromatographic column. *J. Geol.* **95**, 285–307 (1987).
26. McKenzie, D. ($^{230}\text{Th}/^{238}\text{U}$) disequilibrium and the melting process beneath ridge axes. *Earth Planet. Sci. Lett.* **72**, 81–91 (1985).
27. Riter, J. C. A. & Smith, D. Xenolith constraints on the thermal history of the mantle below the Colorado Plateau. *Geology* **24**, 267–270 (1996).
28. Pickett, D. A., Murrell, M. T. & Williams, R. W. Determination of femtomole quantities of protactinium in geologic samples by thermal ionization mass spectrometry. *Anal. Chem.* **66**, 1044–1049 (1994).
29. Edwards, R. L., Cheng, H., Murrell, M. T. & Goldstein, S. J. Protactinium-231 dating of carbonates by thermal ionization mass spectroscopy: implications for the causes of Quaternary climate change. *Science* **276**, 782–786 (1997).

Acknowledgements

We thank V. Salters, C. Lundstrom and D. Pickett for comments on the manuscript and M. Spiegelman, M. Reagan and D.L. Kohlstedt for discussions. We also thank the US Park Service for access. This work was and is being supported by the US National Science Foundation.

Correspondence and requests for materials should be addressed to Y.A. (e-mail: asmerom@unm.edu).

Female bluethroats enhance offspring immunocompetence through extra-pair copulations

Arild Johnsen*, Vegard Andersen, Christine Sunding & Jan T. Lifjeld*

Zoological Museum, University of Oslo, Sars gate 1, N-0562 Oslo, Norway

Female birds frequently copulate with extra-pair males^{1,2}, but the adaptive value of this behaviour is poorly understood². Some studies have suggested that ‘good genes’ may be involved, where females seek to have their eggs fertilized by high-quality males without receiving any material benefits from them^{3,4}. Nevertheless, it remains to be shown that a genetic benefit is passed on to offspring^{5,6}. Here we report that nestling bluethroats, *Luscinia svecica*, sired by extra-pair males had a higher T-cell-mediated

immune response than their maternal half-siblings raised in the same nest. The difference could not be attributed to nestling body mass, sex or hatching order, but may be an effect of paternal genotype. Extra-pair young were also more immunocompetent than their paternal half-sibs raised in the genetic father’s own nest, which indicates an additional effect of maternal genotype. Our results are consistent with the idea that females engage in extra-pair copulations to obtain compatible viability genes, rather than ‘good genes’ *per se*.

Extra-pair mating systems are highly suitable for the study of genetic benefits of mate choice. When eggs in a clutch are fertilized by both the social male and one or more extra-pair males, the effect of paternal genes on offspring fitness can be assessed directly by comparison of maternal half-sibs⁶. This is because the sib groups share the same rearing environment and genes from the mother. The test prediction from ‘good genes’ models is then that extra-pair young (EPY) should perform better than within-pair young (WPY) raised in the same brood. Moreover, if females choose compatible male genes in extra-pair mate choice, it follows that EPY should also perform better than their paternal half-sibs, that is, the WPY of the extra-pair male. Here we have tested these predictions in a population of wild bluethroats where extra-pair paternity occurs at a relatively high frequency^{7,8}.

We studied cell-mediated immunity in nestling bluethroats by a subcutaneous injection of phytohaemagglutinin (PHA) in one

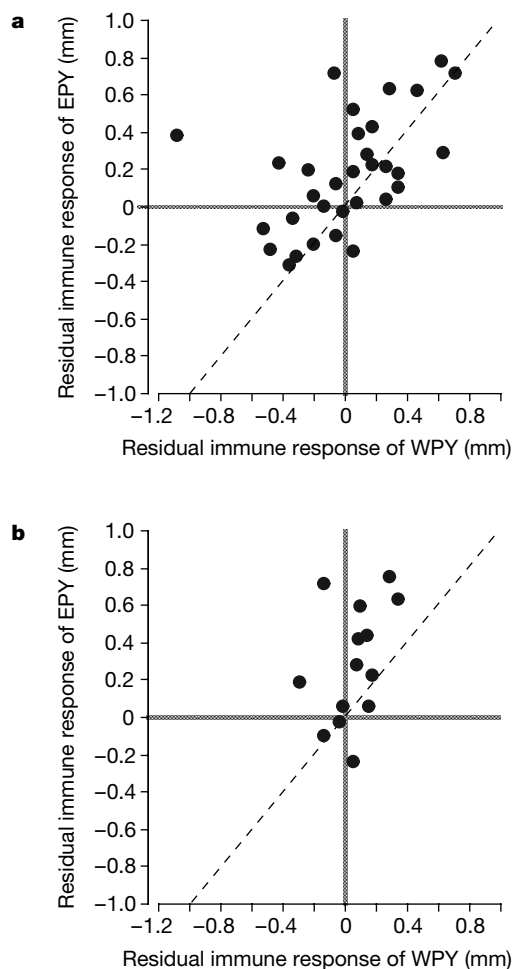


Figure 1 Relationships between residual phytohaemagglutinin (PHA) response of extra-pair young (EPY) and within-pair young (WPY) in bluethroat broods, after controlling for nestling body mass. **a**, Maternal half-sibs raised in the same nest ($n = 32$). **b**, Paternal half-sibs raised in different nests ($n = 14$). Points indicate brood means. Diagonal dashed lines represent identical responses of EPY and WPY.

* Present addresses: Department of Zoology, Göteborg University, PO Box 463, S-405 30 Göteborg, Sweden (A. J.); Department of Biological Sciences, Lapham Hall, PO Box 413, University of Wisconsin Milwaukee, WI 53201, USA (J. T. L.).

wing. PHA causes a local swelling response which peaks after 24 hours and which reflects T-cell reactivity⁹. A genetic component of the PHA response has been documented in cross-foster studies of nestling passerines^{10,11}. It has also been shown that the nestling PHA responses correlate with subsequent survival¹² and longevity¹³.

There was a positive and nearly significant correlation between the PHA response and nestling body mass, using mean values for broods ($r = 0.22$, $n = 75$, $P = 0.063$). An influence of body mass on nestling PHA response has been documented in other passerines^{10,11}. Because we were interested in the genetic component of the PHA response, we used the residuals from a linear regression of wing swelling on body mass for individual nestlings ($r = 0.18$, $n = 372$, $P = 0.0007$) as our measure of immune response.

We found unequivocal support for the prediction that EPY are more immunocompetent than their maternal half-sibs raised in the same nest. EPY and WPY in broods of mixed paternity differed significantly in immune response (Table 1). The response was higher in EPY than in WPY in 24 of the 32 sets of maternal half-sibs (Fig. 1a). The mean difference (EPY – WPY) amounted to 0.51, expressed in phenotypic standard deviations of WPY. EPY and WPY did not differ in body mass growth or tarsus length growth during their first 8 days of the nestling period (Table 1). Hence, genes of extra-pair males appeared to enhance offspring immunocompetence without any corresponding effects on nestling growth.

Although the underlying genetic mechanism for enhanced immunocompetence of EPY is still unknown, a genetic explanation can only be supported through the elimination of potentially confounding, non-genetic factors. For example, EPY might have been able to mount a stronger immune response if they were generally older, that is, they hatched earlier, than their WPY nest mates. This possibility can be rejected as EPY and WPY did not differ in order of hatching rank in broods where the PHA response was measured ($F_{1,30} = 0.28$, $P = 0.60$), nor in all mixed-paternity broods in our sample ($F_{1,37} = 0.02$, $P = 0.89$). Another potential bias might exist if immune responses differed between sexes and the more responsive sex were over-represented among EPY. This idea does not apply either as immune responses did not differ between male and female nestlings in mixed-paternity broods ($F_{1,31} = 0.347$, $P = 0.56$), nor in all mixed-sex broods ($F_{1,62} = 0.138$, $P = 0.71$). The

Table 1 Analyses of variance of immune response and growth in bluethroat nestlings

Test	Effect	Variance component	% of total variance	Mean square	d.f.	F	P
Maternal half-sib comparisons							
Residual PHA response							
	Nest	0.0820	38.5	0.5202	31	4.48	<0.0001
	EPY/WPY	0.0146	6.9	1.0401	1	8.95	0.0033
	Residual	0.1162	54.6	0.1162	126		
Body mass growth							
	Nest	0.0212	55.0	0.1263	32	7.32	<0.0001
	EPY/WPY	0.0001	0.2	0.0232	1	1.34	0.25
	Residual	0.0173	44.8	0.0173	137		
Tarsus length growth							
	Nest	0.0267	76.1	0.1427	23	17.03	<0.0001
	EPY/WPY	0.0000	0.0	0.0073	1	0.87	0.35
	Residual	0.0084	23.9	0.0084	97		
Paternal half-sib comparisons							
Residual PHA response							
	Father	0.0078	4.4	0.1935	13	1.33	0.22
	EPY/WPY	0.0228	12.9	0.9499	1	6.52	0.013
	Residual	0.1457	82.6	0.1457	72		
Body mass growth							
	Father	0.0058	18.5	0.0603	12	1.94	0.13
	EPY/WPY	0.0057	15.2	0.1617	1	5.66	0.032
	Interaction	0.0067	21.4	0.0310	12	2.21	0.021
	Residual	0.0140	45.0	0.0141	58		
Tarsus growth							
	Father	0.0235	45.0	0.1839	8	6.60	<0.0001
	EPY/WPY	0.0008	1.5	0.0469	1	1.68	0.20
	Residual	0.0279	53.4	0.0279	51		

Two-factor analyses of variance, including estimates of variance components, were calculated by the Random Effect option in JMP (version 3.1.1, SAS Institute Inc.). Analyses of maternal half-sibs include only broods with both WPY and EPY, those of paternal half-sibs include only offspring of males siring both WPY and EPY. Calculations of growth rates of body mass and tarsus length are described in Methods. Residual PHA responses were calculated from a linear regression of wing-swelling response on body mass for nestlings in all broods. Sample sizes differ between analyses because of losses owing to predation or because the exact hatching time was unknown.

sex ratio did not differ between EPY and WPY (exact permutation test on within-nest differences: $n = 41$, $P = 0.43$). Differential investment in EPY and WPY by parents also seems unlikely, as no evidence exists that parent birds, including bluethroats¹⁴, can discriminate between their own and extra-pair offspring in a brood¹⁵. There is a theoretical possibility that females can allocate more maternal resources into EPY eggs within a clutch, as female birds can alter the sex ratio¹⁶ or increase testosterone¹⁷ or maternal investments¹⁸ in clutches when socially paired to attractive mates. It remains to be shown, however, that females possess an ability to invest differentially in eggs of different sires within a clutch. If they do, then the only benefit that can select for this behaviour is a difference in offspring reproductive value caused by paternal genes.

We were able to assign paternity to 75% of all EPY in our data set, and 27 of the 84 males had EPY in other males' nests. Identified cuckolders had a similar loss of paternity in their own nest (mean = 23% of brood) to that of males without any EPY (mean = 32% of brood, Mann–Whitney U -test: $Z = -1.32$, $n_1 = 27$, $n_2 = 57$, $P = 0.19$). This is apparently inconsistent with the traditional 'good genes' models, where females are assumed not to differ in mating preferences¹⁹. Furthermore, we have not found any evidence for directional female preferences for male phenotypic traits, as revealed by analyses of male fertilization success in relation to various measures of body size and throat colouration (A.J. and J.T.L., manuscript in preparation). For 14 males, immune responses were measured in both their EPY and WPY. There was a significant effect of paternity on the immune response (Table 1). EPY were more immunocompetent than WPY in 12 of the 14 sets of paternal half-sibs (Fig. 1b). The mean difference (EPY – WPY) amounted to 1.44, expressed in phenotypic standard deviations of the WPY. In these comparisons, it should be noted that EPY and WPY were raised in different nests so that environmental differences might have influenced the results. We found, however, that the immune

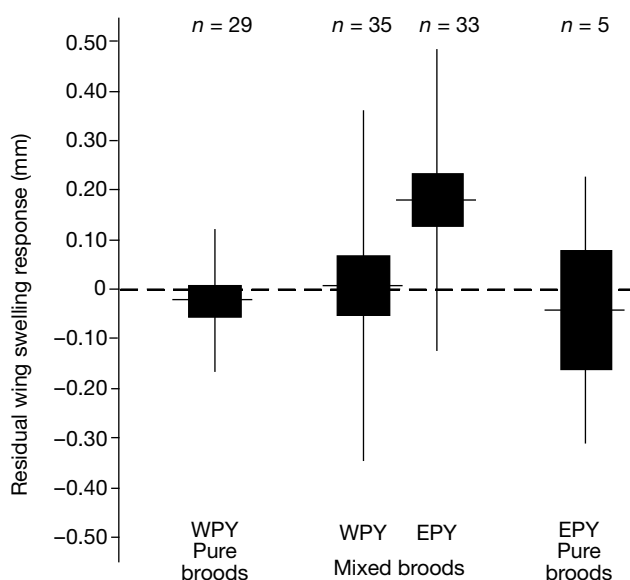


Figure 2 Residual wing-swelling response to phytohaemagglutinin (PHA) injection of nestling bluethroats in relation to brood composition of extra-pair (EPY) and within-pair young (WPY). Horizontal lines indicate brood means, black rectangles indicate ± 1 s.e., and vertical lines indicate ± 1 s.d. In mixed-paternity broods, sample sizes of WPY and EPY differ owing to partial losses of nestlings in either category.

response of the maternal half-sibs of the EPY, that is, the WPY of the cuckolded male, was no higher than that of the WPY of the cuckold (F_{1,11} = 1.70, P = 0.20). Still, EPY were significantly more immunocompetent than their maternal half-sibs in these broods (F_{1,12} = 4.38, P = 0.041). The fact that males produced more immunocompetent offspring with extra-pair females than with their social mates suggests an interaction effect between male and female genotypes; extra-pair mates have a more favourable combination of genes than social mates.

EPY also had a higher body-mass growth rate than their paternal half-sibs raised in the father's own nest (Table 1). For this variable, however, the maternal half-sibs of the EPY, that is, the young sired by the cuckolded male in the same nest, also had a significantly higher growth rate than the WPY in the cuckold's own nest (F_{1,10} = 22.61, P < 0.0001). Thus, the difference in body-mass growth rate between paternal EPY and WPY (Table 1) was probably due to environmental effects rather than genetic effects. The reason behind this pattern is unknown, but it might be explained by males preferentially seeking extra-pair copulations with females in territories of higher quality than their own or by females investing more in their brood after copulation with particular males.

The immune response in broods of only WPY was similar to that of WPY in broods of mixed paternity (Fig. 2: t₆₂ = 0.42, P = 0.67). Most females without any EPY may therefore have failed to obtain a better sire for their offspring. Either they may have been constrained in their extra-pair activities, for example as a consequence of male mate guarding⁷, or compatible males may not have been locally available. In bluethroats, extra-pair sires are usually found among the nearest neighbours^{7,8}. The five broods with only EPY did not have a particularly strong immune response, but it was not significantly lower than that of EPY in mixed-paternity broods (t₃₆ = 1.65, P = 0.11).

Our results from bluethroats provide evidence for a genetic compatibility benefit of extra-pair copulation in female birds. Future work should focus on the relationship between the immune response and offspring fitness, and on the proximate mechanism for how females select extra-pair sires for their offspring. There is an intriguing possibility that females select sperm through a post-copulatory immunological mechanism², which eliminates the need for any phenotypic cues to male genotype. □

Methods

General field procedures

The study was carried out in Øvre Heimdal, S Norway, in 1998 and 1999. A 3-km² study area with roughly 70 territorial males was monitored each year. Details of trapping and blood sampling of adults have been described^{7,8}. A total of 86 broods was included in the study, four of them belonging to two bigamous males. Twelve males and four females were included in the experiment both years. There were no cases of pair mates breeding together twice, or of the same pair of males siring offspring in the same nest twice. Thus, we treated two nesting events by the same adult as independent data. All social fathers were observed providing parental care during the nesting period. Nests were visited frequently around the expected time of hatching, and the hatchlings were weighed to estimate the time of hatching to the nearest hour and determine their order of hatching. Hatchlings were initially marked with a permanent marker and then nail-clipped on day 2 post-hatch for individual recognition. Nestlings were again weighed on day 2, and a blood sample (10–25 µl) was collected. Weighing was also performed on days 5–8 post-hatch in association with the PHA injections. Growth rate of body mass for individual nestlings was expressed as the slope of a linear regression of body mass on the exact age relative to the time of hatching. Nestling tarsus length was measured on day 8, and a growth rate calculated as the length divided by the exact age.

Immune assay

Nestlings received a sensitizing injection of PHA on day 5 post-hatch and a treatment injection on day 7. Each time, 0.10 mg PHA dissolved in 40 µl saline was injected subcutaneously into the right wing. As a control, 40 µl pure saline was injected into the left wing. To avoid confounding swelling effects, the first injection was administered dorsally on the metacarpus (outer) section of the wing and the treatment injection dorsally on the ulna section of the wing. Before injection, the injection site was marked with a permanent marker and the thickness of the wing measured to the nearest 0.01 mm using a spessimeter (Teclock model SM-112). Twenty-four hours (± 1 hour) after injection, the thickness of the wing was measured at the inoculation site. The increase in wing thickness of the right wing minus the change in the left wing was used as an index of the wing-swelling response.

All thickness measurements were done twice, and the repeatability was high, ranging between 0.96 and 0.98. Mean values were used in all calculations. All measurements of a particular brood were performed by one person only. The magnitude of the wing-swelling response more than doubled from the injection on day 5 to that on day 7 (from 0.43 to 0.93 mm, paired t-test on brood means: t₇₀ = 18.26, P < 0.0001). Four broods, which did not receive a sensitizing injection, had a significantly lower swelling response than the rest (t₇₃ = 2.23, P = 0.029) and were therefore excluded from between-brood comparisons, but included in within-brood comparisons.

Paternity and molecular sexing

DNA was extracted from blood samples (in a few cases from embryos in unhatched eggs or dead nestlings) using the QIAamp (Qiagen) extraction kit. Paternity was analysed by polymerase chain reaction (PCR) amplification of six microsatellite loci⁷. The combined exclusion probability²⁰ for these markers was 0.995. The 86 experimental nests contained a total of 480 offspring with sampled DNA. Offspring were defined as WPY if they showed a complete match (n = 335) with their parents, or one paternal mismatch (n = 5) but a low probability of chance inclusion²¹ for the attending male after exclusion of the mismatched locus. Mean probability of chance inclusion for these two groups of nestlings was 0.0056 ± 0.0074 s.d. and 0.0059 ± 0.0076 s.d., respectively. Conversely, offspring were defined as EPY if they showed two or more mismatches (n = 133), or one paternal mismatch (n = 6) but a rather high probability of chance inclusion for the attending male (mean = 0.041 ± 0.038 s.d.) and/or a neighbouring male matching the paternal genotype completely. One offspring showed several mismatches with both parents and was excluded from all analyses (probably an egg dumping by another female). In summary, EPY amounted to 29% (139 out of 479) of all young and occurred in 59% of all broods. Males were assigned paternity to EPY by the same criteria as for social fathers and WPY above. The mean probability of chance inclusion for EPY with complete match with another male (n = 101) was 0.0032 ± 0.0040 s.d., and for EPY with one mismatch (n = 3) the probability was 0.0023, 0.0021 and 0.0069 in each case. One of the microsatellites, Pocc5, shows Z-chromosome linkage in bluethroats⁷. Altogether 380 nestlings could be sexed according to this locus. To sex the remaining nestlings we used the PCR primers P2 and P8 (ref. 22) which amplify two homologous genes: CHD1W on the W chromosome and CHD1Z on the Z chromosome. This marker failed to amplify a product in four cases. We also analysed 41 adults (22 males, 19 females) with this marker. In no case did the Pocc5 or the P2/P8 markers disagree with the phenotypic sex determination of adults based on plumage colouration.

Received 24 March; accepted 9 May 2000.

- Westneat, D. F. & Sherman, P. W. Density and extra-pair fertilizations in birds: a comparative analysis. *Behav. Ecol. Sociobiol.* **41**, 205–215 (1997).
- Birkhead, T. R. In *Sperm Competition and Sexual Selection* (eds Birkhead, T. R. & Møller, A. P.) 579–622 (Academic, London, 1998).
- Kempnaers, B. et al. Extra-pair paternity results from female preference for high-quality males in the blue tit. *Nature* **357**, 494–496 (1992).
- Hasselquist, D., Bensch, S. & von Schantz, T. Correlation between male song repertoire, extra-pair paternity and offspring survival in the great reed warbler. *Nature* **381**, 229–232 (1996).
- Petrie, M. Improved growth and survival of offspring of peacocks with more elaborate trains. *Nature* **327**, 598–599 (1994).
- Sheldon, B. C., Merilä, J., Qvarnström, A., Gustafsson, L. & Ellegren, H. Paternal genetic contribution to offspring condition predicted by size of male secondary sexual character. *Proc. R. Soc. Lond. B* **264**, 297–302 (1997).
- Johnsen, A., Lifjeld, J. T., Rohde, P. A., Primmer, C. R. & Ellegren, H. Sexual conflict over fertilizations: female bluethroats escape male paternity guards. *Behav. Ecol. Sociobiol.* **43**, 401–408 (1998).
- Johnsen, A., Andersson, S., Örnberg, J. & Lifjeld, J. T. Ultraviolet plumage ornamentation affects social mate choice and sperm competition in bluethroats (*Aves: Luscinia s. svecica*): a field experiment. *Proc. R. Soc. Lond. B* **265**, 1313–1318 (1998).
- McCorkle, F. Jr, Olah, I. & Glick, B. The morphology of the phytohemagglutinin-induced cell response in the chicken wattle. *Poultry Sci.* **59**, 616–623 (1980).
- Saino, N., Calza, S. & Møller, A. P. Immunocompetence of nestling barn swallows in relation to brood size and parental effort. *J. Anim. Ecol.* **66**, 827–836 (1997).
- Brinkhof, M. W. G., Heeb, P., Kölliker, M. & Richner, H. Immunocompetence of nestling great tits in relation to rearing environment and parentage. *Proc. R. Soc. Lond. B* **266**, 2315–2322 (1999).
- Christe, P., Møller, A. P. & de Lope, F. Immunocompetence and nestling survival in the house martin: the tasty chick hypothesis. *Oikos* **83**, 175–179 (1998).
- Birkhead, T. R., Fletcher, F. & Pellatt, E. J. Nestling diet, secondary sexual traits and fitness in the zebra finch. *Proc. R. Soc. Lond. B* **266**, 385–390 (1999).
- Johnsen, A., Krokene, C. & Lifjeld, J. T. Brood division is associated with fledgling dispersion in the bluethroat (*Luscinia s. svecica*). *Auk* **114**, 553–561 (1997).
- Kempnaers, B. & Sheldon, B. C. Why do male birds not discriminate between their own and extra-pair offspring? *Anim. Behav.* **51**, 1165–1173 (1996).
- Sheldon, B. C., Andersson, S., Griffith, S. C., Örnberg, J. & Sendecka, J. Ultraviolet colour variation influences blue tit sex ratios. *Nature* **402**, 874–877 (1999).
- Gil, D., Graves, J., Hazon, N. & Wells, A. Male attractiveness and differential testosterone investment in zebra finch eggs. *Science* **286**, 126–128 (1999).
- Cunningham, E. J. A. & Russell, A. F. Egg investment is influenced by male attractiveness in the mallard. *Nature* **404**, 74–77 (2000).
- Andersson, M. *Sexual Selection* (Princeton Univ. Press, Princeton, 1994).
- Jamieson, A. The effectiveness of using co-dominant polymorphic allelic series for (1) checking pedigrees and (2) distinguishing full-sib pair members. *Anim. Genet.* **25**, 37–44 (1994).
- Jeffreys, A. J., Allen, M. J., Hagelberg, E. & Sonberg, A. Identification of the skeletal remains of Josef Mengele by DNA analysis. *Forensic Sci. Int.* **56**, 65–76 (1992).
- Griffiths, R., Double, M. C., Orr, K. & Dawson, R. J. G. A DNA test to sex most birds. *Mol. Ecol.* **7**, 1071–1075 (1998).

Acknowledgements

We thank B.-A. Bjerke and C. K. Aas for assistance, and T. Amundsen, M. Andersson, T. R. Birkhead, D. Blomqvist, L. T. T. Hansen, B. Kempnaers, F. Mehlum, M. Petrie and T. Slagsvold, for comments. This work was supported by a post-doc fellowship from the Norwegian Research Council (A.J.) and a grant from the Nansen Endowment (J.T.L.). Permission to carry out the immune experiment on nestling bluethroats was given by the Norwegian Animal Research Authority.

Correspondence and requests for materials should be addressed to J.T.L. (e-mail: j.t.lifjeld@toyen.uio.no).

Magnetite defines a vertebrate magnetoreceptor

Carol E. Diebel^{††}, Roger Proksch^{‡§}, Colin R. Green^{||}, Peter Neilson[‡] & Michael M. Walker^{*}

^{*} Experimental Biology Research Group, School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland, New Zealand

[‡] Magnetism Laboratory, Digital Instruments, Santa Barbara, California 93117, USA

^{||} Department of Anatomy with Radiology, School of Medicine, University of Auckland, Private Bag 92019, Auckland, New Zealand

The key behavioural, physiological and anatomical components of a magnetite-based magnetic sense have been demonstrated in rainbow trout (*Oncorhynchus mykiss*)¹. Candidate receptor cells located within a discrete sub-layer of the olfactory lamellae contained iron-rich crystals that were similar in size and shape to magnetite crystals extracted from salmon^{1,2}. Here we show that these crystals, which mapped to individual receptors using confocal and atomic force microscopy, are magnetic, as they are uniquely associated with dipoles detected by magnetic force microscopy. Analysis of their magnetic properties identifies the crystals as single-domain magnetite. In addition, three-dimensional reconstruction of the candidate receptors using confocal and atomic force microscopy imaging confirm that several magnetic crystals are arranged in a chain of about 1 μm within the receptor, and that the receptor is a multi-lobed single cell. These results are consistent with a magnetite-based detection mechanism^{2,3}, as 1- μm chains of single-domain magnetite crystals are highly suitable for the behavioural and physiological responses to magnetic intensity previously reported in the trout.

Permanently magnetized single-domain magnetite crystals ('lodestone', $\text{Fe-O Fe}_2\text{O}_3$) have been described in many different phyla², and the suitability of magnetite for magnetoreception has been analysed in detail²⁻⁵. If the crystals in the nose of the trout are single-domain magnetite, they could act as detector elements within a receptor if their movement, or the torque on them, arising from the interaction of their magnetic moments with the earth's magnetic field is measurable by the nervous system⁶. Single 50-nm crystals of magnetite do not interact strongly enough with the earth's magnetic field to overcome the randomizing effects of thermal buffeting²; however, if the crystals are arranged in chains as in the magnetotactic bacteria, their individual moments will sum linearly. The average orientation of a freely rotating chain will be aligned with the external field vector, whereas the variance of the chain's orientation will depend on the intensity of the external field. Chains of crystals with magnetic-to-thermal energy ratios of two and six are optimal for detecting magnetic intensity and direction,

respectively². A consistent chain length giving either of the above ratios will thus provide evidence of selection for use of the chains in magnetoreception.

The existence of magnetite-based receptors is difficult to demonstrate because the magnetite crystals are too small to be detected easily and the receptors do not need to be aggregated into a complex sense organ^{1,7}. We previously imaged putative magnetite crystals in thin sections from the nose of the trout but were unable to identify the crystals uniquely nor show that the crystals were organized for use in magnetoreception¹. To overcome the problems of scale presented by the small size and low volume concentrations of the crystals, we have now combined confocal laser scanning microscopy (CLSM) with atomic force microscopy (AFM)/magnetic force microscopy (MFM) techniques to characterize the putative magnetite crystals and the candidate magnetoreceptor cells in the trout.

Magnetic force microscopy is capable of determining magnetic domain structure in a variety of magnetic materials, including small particles with a spatial resolution of less than 100 nm. Because it is

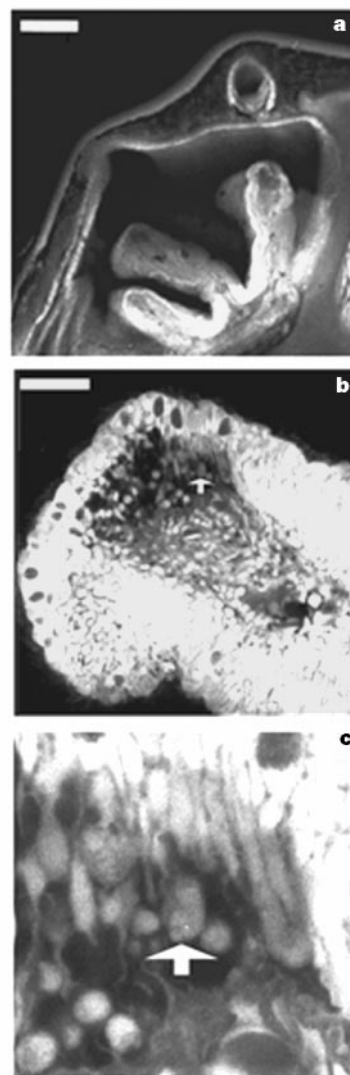


Figure 1 Step-wise increase in the magnification (CLSM autofluorescent images) of the area in the olfactory lamellae where we find magnetic crystals. **a**, Section through a trout nasal capsule with three lamellae evident within the capsule; a lateral line canal (llc) is also seen. Scale bar, 250 μm . **b**, Section through the distal half of the middle lamella in **a**, with an arrowhead pointing to the magnetoreceptor cell shown in **c**. Scale bar, 50 μm . **c**, CLSM overlay (60- μm square) taken around the magnetoreceptor cell shown in **b**. The image of the reflectance contained in the cell has been overlaid onto the autofluorescent image of the cell.

Present addresses: [†] Auckland Museum, Private Bag 92018, Auckland, New Zealand and [§] Asylum Research, 601 Pine Avenue, Santa Barbara, California 93117, USA.

sensitive to magnetic forces, it can also image magnetic structures that are covered by a layer of non-magnetic material. Both calculations and experimental results from magnetotactic bacteria have shown that MFM has sufficient sensitivity to image single-domain magnetite particles⁸. Experimentally, it has been observed that the MFM shows magnetic contrast above single-domain magnetite particles when the MFM tip is within 1 μm of the particle. To prepare the trout tissue for AFM/MFM imaging, we sectioned JB4 blocks of tissue along the z -axis until candidate particles located with the CLSM were within 100–200 nm of the block surface (Fig. 1). The location of the candidate particles was then reconfirmed and image maps generated using the CLSM to assist in precise positioning of the scanning tip of the atomic/magnetic force microscope to the same locale.

Subsequent MFM only detected a distinct 'magnetic contrast' at the locations where reflectance had previously been observed in CLSM (Fig. 2a, b). The areas showing magnetic contrast were less than 100 nm in diameter and had dipole moments of about 10^{-17} to 10^{-16} Ampere-metre² (A-m^2) (Figs 2c and 3). Although the MFM image did not indicate whether the source of the magnetic contrast is a single large magnetic particle or several smaller particles, simultaneous AFM topography images showed that the magnetic sources arose from small, uniformly sized particles present either singly or in small groups (Fig. 2a).

When we applied a magnetic field, the magnetization of the particle(s) reversed. Observation of this sort of reversal in the presence of a non-zero applied magnetic field is a clear indication of magnetic hysteresis associated with ferromagnetic samples⁹. This 'smoking gun' (changes in polarity in a non-zero applied field) is visible in the four images of Fig. 3. The first image (Fig. 3a) shows a dark patch in the region of the magnetic source. This is consistent with an attractive interaction caused by the magnetic field from the MFM tip magnetizing the particle. The next image (Fig. 3b) was made in an applied field of +130 millitesla (mT) and shows a dipolar contrast (black/white), consistent with the magnetite particle being magnetized in the plane of the sample. This field value is below the 500-mT coercivity of the MFM tip that we used, insuring that the tip magnetization remained primarily along the axis of the tip. The two subsequent images (Fig. 3c, d) show the polarity of the imaged dipole switching in response to toggling the applied field.

We calculated the MFM image switching-field distribution

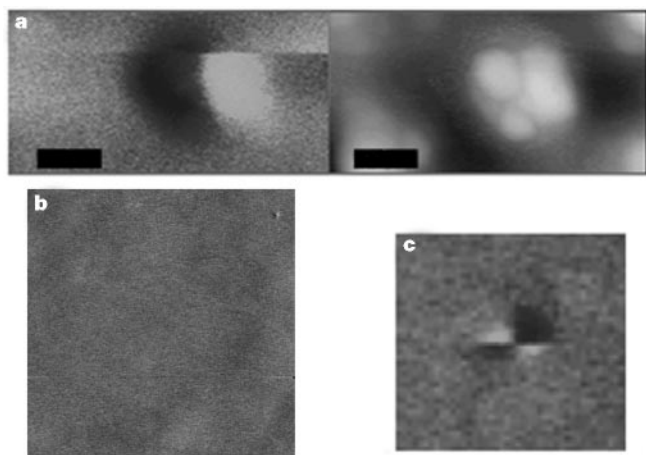


Figure 2 Images of magnetic particle(s). **a**, Scanned images of multiple particles taken simultaneously in MFM (left) and AFM (right) mode. The dipole image shown on the left is the summed field of the three particles shown on the right. Scale bar, 100 nm. **b**, MFM scan taken of a 10- μm square area in a lamella. Note the small dipole present in the upper right hand corner. **c**, MFM image (75-nm square) of a magnetic dipole switching because of the applied field of +25 mT and the field from the MFM tip.

(Fig. 4) for a particle by subtracting sequential images¹⁰. We measured a coercivity between 20 and 40 mT for the particle(s), which is consistent with measurements made on single-domain magnetite particles from magnetotactic bacteria^{10,11}. The magnetic field from the MFM tip will affect this value and therefore it must be viewed with some caution^{9,10}. Magnetite in sockeye salmon had a mean coercivity of roughly 30 mT (ref. 12), whereas the mean moment of magnetite particles extracted from the salmon was about 2×10^{-17} A-m^2 (ref. 13). These values both agree well with observed values for the single and grouped particles shown in Figs 2 and 3 and also exclude greigite as the magnetic source. The coercivity and magnetic moment of greigite particles of the sizes found in the trout would be less than 2.5 mT and 0.5×10^{-17} A-m^2 respectively¹⁴. In addition, we took two of the same MFM tips used on the crystals within the trout tissue and imaged magnetotactic bacteria samples. This provided a calibration sample, as the moment of the magnetite crystals within the bacteria is well defined. The calculated moment of the group of multiple crystals shown in Fig. 2 in the trout tissue using this technique is 5×10^{-16} A-m^2 , which is consistent with several single-domain-sized particles being the source of contrast in the trout images.

The reflectance of the crystals obtained on the CLSM can be mapped to a multi-lobed cell that is located within the basal lamina of a trout olfactory lamella (Fig. 5). The cell is 10–15 μm in diameter and has several processes (>4) that extend out to and are surrounded by tubular-shaped fibroblastic cells (with two processes) that help delineate the basal layer (Fig. 5a, b). Analysis of the 3D images taken in both fluorescent and reflectance CLSM mode of at least four cells gave an estimated length of the crystal chain of about 1 μm (Fig. 5c, d). The observation of single particles in AFM that are associated with the smallest magnetic sources detected in MFM thus seems likely to occur when a chain is lying end on to the surface of the block (for example, Fig. 3). Conversely, groups of particles might be observed either where the chain of

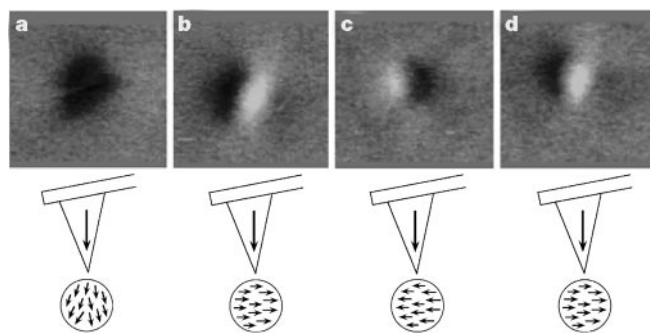


Figure 3 MFM images that show the response of a putative single magnetic particle (within trout tissue) in the presence of an applied field. The magnetic field applied in the plane of the sample was +1.4, +150, -150 and +130 mT for images **a–d**, respectively. MFM images (75-nm squares) are shown on top, with a representation of the MFM tip and magnetization of the particle underneath. The MFM tip (inverted triangle) is permanently magnetized with a coercivity of +500 mT at right angles (arrow in inverted triangle) to the applied field. The small arrows within each circle under the tip represent the alignment of the individual magnetic dipole moments that might act as the field source. **a**, Image shows a dark patch at the location of the particle. This dark patch indicates an attractive reaction between the tip and sample, consistent with the magnetic field from the MFM tip weakly magnetizing the particle and causing an attractive interaction. **b–d**, MFM images show the nearly dipolar responses of the magnetic particle under a strong applied magnetic field. These are consistent with an MFM image of a single-domain particle magnetized along the direction of the applied field. Note that the reversal of the field and dipolar response in **c** are consistent with the particle magnetization flipping in the reversed applied field. In images **b–d**, The applied field was large enough to completely align the magnetic moment of the crystals within the field.

particles bends (evident in Fig. 5c and d) or where strong attractions between opposite poles may occasionally cause aggregation of the crystals during fixation and embedding of the tissue samples.

The distinctive shape, consistent co-localization within the same sub-layer where the magnetically responsive branch of the trigeminal nerve terminates, and the presence of single-domain magnetite crystals lead us to the conclusion that these cells are excellent

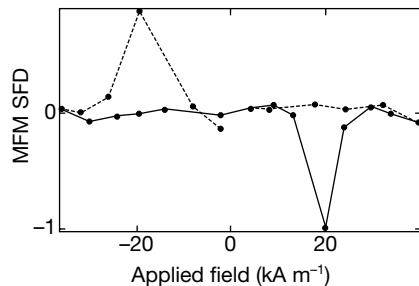


Figure 4 MFM image switching-field distribution. The switching-field distribution was measured by subtraction of sequential images of the MFM signal above the putative dipole as the applied field was systematically varied. The results were scaled to give a maximum magnitude of 1. There are two large spikes, one corresponding to a positive field value, another to a negative. The separation between the two peaks is about 40 kA m^{-1} . This implies a coercivity of 20 kA m^{-1} , consistent with values found in bacterial magnetite, also measured with an MFM¹⁰.

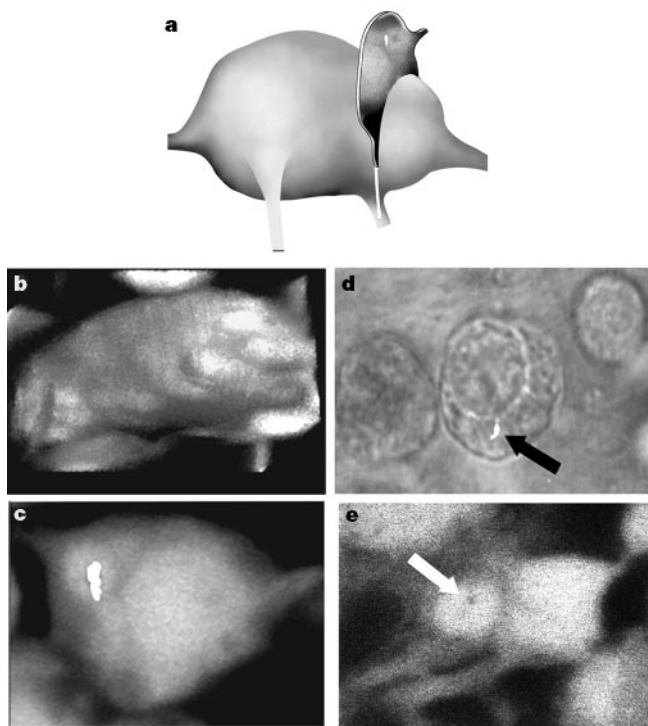


Figure 5 Images of the magnetoreceptor cell. **a**, Representation of a magnetoreceptor. A single optical slice (also shown in **c**) taken at the level of the magnetic particles is extended out of the receptor to show the placement of the magnetic particle chain. **b**, Three-dimensional volume rendered image of a single magnetoreceptor cell that shows its multi-lobed shape. **c**, CLSM optical slice overlay of an autofluorescent image of receptor and reflectance image of a $1\text{-}\mu\text{m}$ chain of magnetic particles (arrow). **d**, CLSM $30\text{-}\mu\text{m}$ wide overlay in reflected and transmitted light modes of a magnetoreceptor isolated from a trypsin digest of an olfactory lamella. The arrow indicates a chain of reflecting crystals. **e**, CLSM optical slice ($\sim 30\text{-}\mu\text{m}$ wide) taken through a magnetoreceptor within an olfactory lamella. The arrow points to the reflectance of a particle contained within the cell.

candidates for magnetoreceptors. We note that the moments for the magnetite particles extracted from the salmon suggest the chains of crystals will have a magnetic to thermal energy ratio of about four, which is consistent with their use in both the behavioural and physiological responses to magnetic fields that have been shown in a variety of taxa^{15–18}. This discovery of magnetite in a form that appears optimal for detecting changes in magnetic fields raises the possibility that magnetite is the basis of a general magnetoreceptor mechanism in taxa as distant as the vertebrates¹⁵ and arthropods^{17,19,20}. We also note that the magnetite discovered in the trout would be suitable for detecting the small changes in magnetic intensity required by a new model of magnetic position determination by homing pigeons^{21,22}. We therefore suggest that these cells are very likely to be the site of magnetic field detection in the trout and that understanding should now be sought of how the chains of crystals could transduce a magnetic field into an electrical signal in the nervous system. □

Methods

Tissue preparation

Juvenile trout (40–60 mm) were anaesthetized in 0.05% MS 222, measured and decapitated just behind the gill cover. After immersion in fixative for 24 h, the head was embedded in JB4 resin. Thin sections were taken through the head to expose the olfactory lamellae for microscopic analysis.

Confocal microscopy

Candidate magnetic crystals were located and mapped within the lamellae using reflectance-mode on a Leica TCS 4D confocal laser scanning microscope. A combined filter set (488/568) was then used to image and map the area around the candidate crystals using tissue and glutaraldehyde induced autofluorescence. The reflecting crystals were found within single cells, and three-dimensional rendering based upon 50-nm optical slices (in both autofluorescent and reflectance mode) was done using VoxelfView (Vital Images, Fairfield, Iowa) to map precisely the location of the magnetite crystal chain within a cell.

Atomic/magnetic force microscopy

A commercially available atomic/magnetic force microscope (Dimension 3100, Digital Instruments, Santa Barbara, CA) was used to generate all the images. The microscope was operated in Tapping/Lift mode, a well established technique for imaging both topography and magnetic force over a region of the sample⁹. This mode provides simultaneous high-resolution AFM images of the sample surface topography and magnetic interactions made at a well defined tip–sample separation referred to as the ‘lift height’. The imaging was done in the presence of an applied magnetic field, using a rotating rare earth magnet²³, and a high coercivity tip²⁴. This had two primary benefits: first, the applied field aligned the magnetic moment of the sample leading to stronger MFM contrast; and second, the direction of the applied field could be reversed, in turn reversing the magnetic moment of the sample and providing an unambiguous indication of the existence of magnetism in the trout tissue.

Received 3 December 1999; accepted 7 April 2000.

- Walker, M. M. *et al.* Structure and function of the vertebrate magnetic sense. *Nature* **390**, 371–376 (1997).
- Kirschvink, J. L. & Walker, M. M. in *Magnetite Biomineralization and Magnetoreception by Living Organisms: A New Biomagnetism* (eds Kirschvink, J. L., Jones, D. S. & MacFadden, B. J.) 243–254 (Plenum, New York, 1985).
- Kirschvink, J. L. & Gould, J. L. Biogenic magnetite as a basis for magnetic field detection in animals. *Biosystems* **13**, 181–201 (1981).
- Yorke, E. D. A possible magnetic transducer in birds. *J. Theor. Biol.* **77**, 101–105 (1979).
- Yorke, E. D. Sensitivity of pigeons to small magnetic field variations. *J. theor. Biol.* **89**, 533–537 (1981).
- Gould, J. L. Magnetic field sensitivity in animals. *Ann. Rev. Physiol.* **46**, 585–598 (1984).
- Kirschvink, J. L. Biogenic ferrimagnetism: A new biomagnetism, in *Biomagnetism* (eds Williamson, S. J., Romani, G. L., Kaufman, L. & Modena, I.) 501–531 (Plenum, New York, 1983).
- Prokisch, R. B. *et al.* Magnetic force microscopy of the submicron magnetic assembly in a magnetotactic bacterium. *Appl. Phys. Lett.* **66**, 2582–2584 (1995).
- Babcock, K., Dugas, M., Manalis, S. & Elings, V. Magnetic force microscopy. *Res. Soc. Symp. Proc.* **355**, 311 (1995).
- Wittborn, J. *et al.* Magnetization reversal observation and manipulation of chains of nanoscale magnetic particles using the magnetic force microscope. *Nano. Mater.* **12**, 1149–1152 (1999).
- Dunin-Borkowski, R. E. *et al.* Magnetic microstructure of magnetotactic bacteria by electron holography. *Science* **282**, 1868–1870 (1998).
- Walker, M. M., Quinn, T. P., Kirschvink, J. L. & Groot, C. Production of single-domain magnetite throughout life by sockeye salmon, *Oncorhynchus nerka*. *J. Exp. Biol.* **140**, 51–63 (1988).
- Mann, S., Sparks, N. H. C., Walker, M. M. & Kirschvink, J. L. Ultrastructure, morphology and organization of biogenic magnetite from sockeye salmon, *Oncorhynchus nerka*; implications for magnetoreception. *J. Exp. Biol.* **140**, 35–49 (1988).
- Diaz-Ricci, J. C. & Kirschvink, J. L. Magnetic domain state and coercivity predictions for biogenic greigite (Fe_3S_4): A comparison of theory with magnetosome observations. *J. Geophys. Res.* **97**, 17039–

- 17315 (1992).
15. Semm, P. & Beason, R. C. Responses to small magnetic field variations by the trigeminal system of the bobolink. *Brain Res. Bull.* **25**, 735–740 (1990).
 16. Walker, M. M. Learned magnetic field discrimination in the yellowfin tuna, *Thunnus albacares*. *J. Comp. Physiol. A* **155**, 673–679 (1984).
 17. Walker, M. M. & Bitterman, M. E. Honeybees can be trained to respond to very small changes in geomagnetic field intensity. *J. Exp. Biol.* **145**, 489–494 (1989).
 18. Wiltschko, R. & Wiltschko, W. *Magnetic Orientation in Animals*. (Springer, Berlin, Heidelberg, New York, 1995).
 19. Gould, J. L., Kirschvink, J. L., Deffeyes, K. S. Bees have magnetic remanence. *Science* **201**, 1026–1028 (1978).
 20. Lohmann, K. J. Magnetic remanence in the Western Atlantic spiny lobster. *J. Exp. Biol.* **113**, 29–41 (1984).
 21. Walker, M. M. On a wing and a vector: A model for magnetic navigation by homing pigeons. *J. Theor. Biol.* **192**, 341–349 (1998).
 22. Walker, M. M. Magnetic position determination by homing pigeons. *J. Theor. Biol.* **197**, 271–276 (1999).
 23. Proksch, R. B., Runge, E., Hansma, P. K., Foss, S. & Walsh, B. High field magnetic force microscopy. *J. Appl. Phys.* **78**, 3303–3307 (1995).
 24. Liou, S. H. & Yao, Y. D. Development of high coercivity magnetic force microscopy tips. *J. Magn. Mater.* **190**, 130–134 (1998).

Acknowledgements

We thank K. Babcock at Digital Imaging for the generous use of the AFM/MFM. The Biological Imaging Research Unit at the School of Medicine, University of Auckland, provided the CLSM and imaging facilities. In addition, we thank S. Edgar, A. Turner, H. Holloway and especially B. Beaumont for their assistance in preparation and viewing samples on the CLSM and transmission electron microscopy. Financial support came from the Marsden Fund and the School of Biological Sciences.

Correspondence and requests for materials should be addressed to C.E.D. (e-mail: cdiebel@akmuseum.org.nz) or M.M.W. (e-mail: m.walker@auckland.ac.nz).

Transformation from temporal to rate coding in a somatosensory thalamocortical pathway

Ehud Ahissar, Ronen Sosnik & Sebastian Haidarliu

Department of Neurobiology, The Weizmann Institute of Science, Rehovot 76100, Israel

The anatomical connections from the whiskers to the rodent somatosensory (barrel) cortex form two parallel (lemniscal and paralemniscal) pathways^{1,2}. It is unclear whether the paralemniscal pathway is directly involved in tactile processing, because paralemniscal neuronal responses show poor spatial resolution, labile latencies and strong dependence on cortical feedback^{3–5}. Here we show that the paralemniscal system can transform temporally encoded vibrissal information into a rate code. We recorded the representations of the frequency of whisker movement along the two pathways in anaesthetized rats. In response to varying stimulus frequencies, the lemniscal neurons exhibited amplitude modulations and constant latencies. In contrast, paralemniscal neurons in both thalamus and cortex coded the input frequency as changes in latency. Because the onset latencies increased and the offset latencies remained constant, the latency increments were translated into a rate code: increasing onset latencies led to lower spike counts. A thalamocortical loop that includes cortical oscillations and thalamic gating can account for these results. Thus, variable latencies and effective cortical feedback in the paralemniscal system can serve the processing of temporal sensory cues, such as those that encode object location during whisking. In contrast, fixed time locking in the lemniscal system is crucial for reliable spatial processing.

The lemniscal pathway of the rat trigeminal system ascends

through the ventral posterior medial nucleus (VPM) of the thalamus to the barrels in layer 4 of the cortex and to layers 5b and 6 (refs 6, 7). The paralemniscal pathway ascends through the medial division of the posterior nucleus (POm) of the thalamus to layers 1 and 5a and to the septa between the barrels in layer 4 (refs 7, 8). To reveal the processing differences between the two pathways, we recorded from single-units ($n = 710$) and multi-units ($n = 540$) of the major stations along these pathways while we stimulated (moved) the whiskers. First we analysed single- and multi-units separately. Analysis of the single-units revealed that response patterns were usually similar for neighbouring neurons. Therefore, spikes of all single- and multi-units that were recorded simultaneously from a single electrode were pooled, and these pools were referred to as 'local populations'.

First, we examined responses to stimuli that mimic natural whisking conditions⁹: 50-ms pulses of air puffs applied to one or two rows of whiskers at 8 Hz. Typical recordings along both pathways are shown in Fig. 1. Brainstem neurons appeared simply to

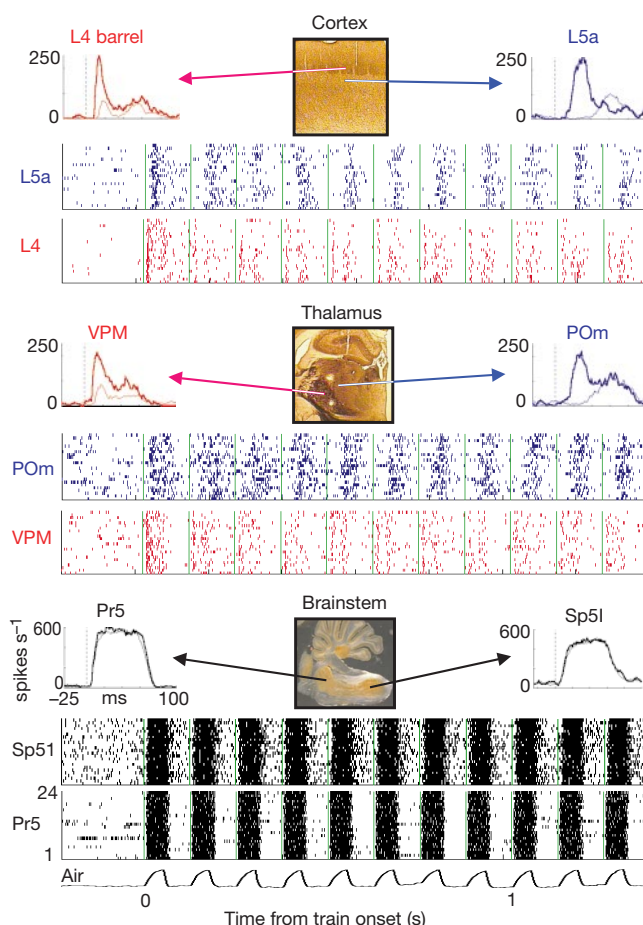


Figure 1 Lemniscal and paralemniscal thalamic transformations. Recordings of local populations from the six major stations along both pathways. Raster displays show firing times of recorded neurons (black, red or blue vertical lines) in relation to stimulus onset times (green vertical lines). Responses to 24 stimulus trains are plotted. PSTHs were computed for each stimulus cycle and are shown for the first (at time 0; thick curves) and last three cycles of the figure (cycles 9–11; thin curves). Response patterns remained steady during the remaining 13 cycles (not shown). The main trigeminal nuclei (Pr5 and Sp5l) receive vibrissal input and project to both thalamic nuclei, with a bias of Pr5 towards VPM and Sp5l towards POm^{16,29}. VPM projects to cortical barrels (dark areas in layer 4) and layers 5b and 6, and receives feedback from the upper part of layer 6 (refs 6, 7, 30). POm projects to layers 1 and 5a, and the septa between the barrels in layer 4 (refs 7, 8), and receives feedback from layer 5 and the lower part of layer 6 (refs 7, 30). Electrolytic lesions in the thalamus mark recording locations of single local populations in the POm (upper) and VPM.

relay primary afferent responses¹⁰: the response to a single pulse was constant along the train, as evident from the raster display and from the overlap between the peristimulus time histograms (PSTHs) computed for the first cycle (thick lines) and for the last three cycles (thin lines). However, as expected^{4,11}, thalamic neurons did not behave as relays, and their outputs exhibited specific transformations of the brainstem signals (Fig. 1, middle rows). The transformation at the VPM resembled an adaptation process during which, following the first stimulation cycle, the spike count (per stimulus cycle) decreased as the train continued, until a 'steady state' (a period during which the spike count was steady) was reached. Recordings from the POm revealed a different stabilization process. The latency of the response onset increased during the 8-Hz train until it stabilized at a significantly longer steady-state value (Fig. 1, POm raster display and PSTHs). These latency shifts resulted in decreased spike counts in the POm neurons because offset latencies did not change. In the barrel cortex there was a similar dissociation for the response patterns of the lemniscal and paralemniscal projection targets (Fig. 1, top rows). Whereas the responses of layer 4 barrel neurons were similar to those of VPM neurons, the responses of layer 5a neurons resembled those of POm neurons. These response patterns were so characteristic for single- and multi-units at each station that the recording location could often be predicted based on the physiological responses, even before histological processing.

Next, we tested the dependency of the thalamic transformations on the frequency of whisker movement using three frequencies (2, 5 and 8 Hz), revealing another dissociation between the two pathways. Whereas steady-state latencies of the lemniscal pathway were almost constant for all the tested frequencies, those of the paralemniscal pathway increased with increasing frequencies. The path-

way-specific response patterns are demonstrated by the PSTHs that were averaged, for each station, across all neurons recorded from well localized sites (Fig. 2). At a stimulation frequency of 2 Hz (below the natural whisking range), latencies at all stations barely changed between the first stimulus cycles (left insets) and the steady-state periods (centre PSTHs). With stimulation frequencies of 5 and 8 Hz (both of which are within the whisking frequency range¹²), brainstem and lemniscal latencies were also essentially constant (Fig. 2, centre PSTHs and right insets), in agreement with other findings¹⁰. However, steady-state latencies of the paralemniscal pathway increased with increasing stimulus frequency (Fig. 2, centre PSTHs and right insets; see also ref. 3).

For both pathways, the magnitude of the steady-state responses (measured by the PSTH area, that is, spike counts per cycle) decreased with increasing frequencies. Again, in the lemniscal pathway, spike-count decrements were caused by amplitude reduction, whereas in the paralemniscal pathway they resulted primarily from latency shifts. As a result, during steady-state periods of whisker stimulation, whisker frequencies were represented by spike counts (rate coding) in both pathways, and by latencies (temporal coding) in the paralemniscal pathway. These representations were reproduced during each stimulus train, as the responses to the first cycle in each train were identical for all frequencies (Fig. 2, left insets). The averaged latency and spike-count representations of the whisker frequency (Fig. 2, right insets) depict typical representations of single local populations in each of the stations. These pathway-specific representations occurred for all well localized recording sites, except for 1 out of 14 in POm and 1 out of 13 in layer 5a, in both of which the latency increased between 2 and 5 Hz but not between 5 and 8 Hz. Part of the amplitude adaptation could be attributed to anaesthesia¹⁰. However, anaesthesia could not account for the latency shifts in the paralemniscal system, because latency shifts due to anaesthesia are an order of magnitude smaller^{13–15}. Furthermore, as these latency shifts developed during each stimulus train, they probably reflect a dynamic process.

Paralemniscal processing could be conducted in parallel to, or in series with^{2,4,5}, lemniscal processing. If processing is in series, the onset of POm activity should lag the onset of lemniscal cortical activity. However, as seen in Fig. 3, which shows the average responses to the first cycle of mechanical stimulation applied to the principal whiskers, the average POm response rises before those of layers 4 and 5b. Furthermore, the shortest and modal latencies to the first stimulus cycle in the POm (7 and 8 ms, respectively) were shorter than those in layer 4 (8 and 9 ms) and layer 5b (8 and 9 ms). Thus, the onset of POm activity must be driven directly by the onset

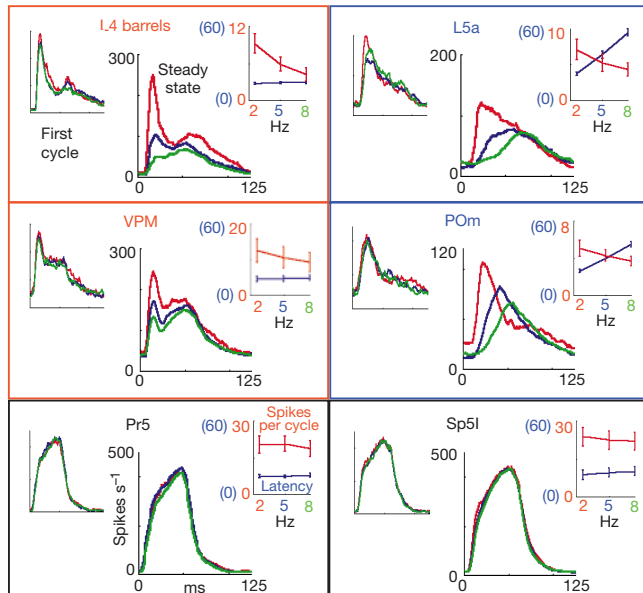


Figure 2 Response transformations and steady-state frequency representations. For each station, data was averaged across all well localized local populations (5, 5, 7, 14, 9 and 13 local populations from Pr5, Sp5l, VPM, POm, layer 4 barrels and layer 5a, respectively). Air puffs (50 ms) moved the whiskers anteriorly (protraction direction) at 2 (red), 5 (blue) and 8 Hz (green). For each station, the average PSTH for the first stimulus cycle (left) and steady-state period (0.5–3 s; middle; scales as for first stimulus cycle) are shown. On the right of each panel are the average ($\pm$ s.e.m.) spike count (red) and latency (blue) as a function of stimulation frequency. These averages were computed across all well localized local populations of each station during steady states. Latencies to half peak are shown. A significant dependency of latency on frequency was found only in POm and layer 5a (POm, $r^2 = 0.62$, $P < 0.0001$; layer 5a, $r^2 = 0.71$, $P < 0.0001$; in all brainstem and lemniscal stations, slope $< 0.5 \text{ ms Hz}^{-1}$, $r^2 < 0.03$ and $P > 0.44$).

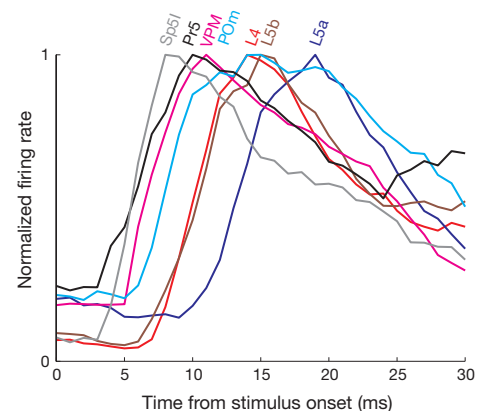


Figure 3 Order of activation onset. Average PSTHs of the entire population recorded at each station during the first cycle of a mechanical stimulus train applied to the principal whisker (forward deflections using the mechanical stimulator). Sp5l ($n = 14$ units), Pr5 ($n = 10$), VPM ($n = 12$), POm ($n = 17$), layer 4 barrels (L4; $n = 22$), layer 5b (L5b; $n = 24$) and layer 5a (L5a; $n = 26$).

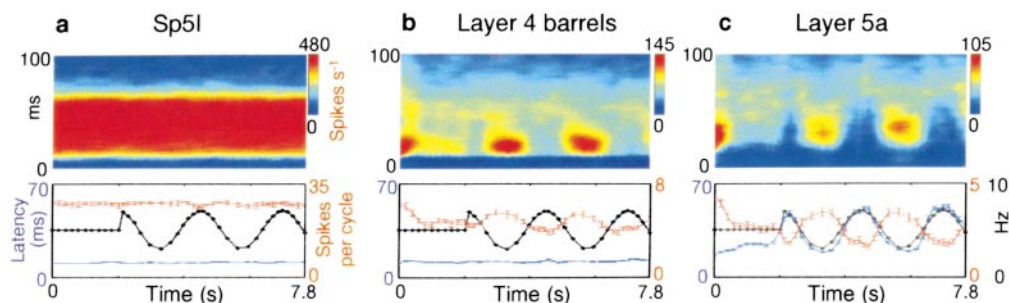


Figure 4 Representations of the instantaneous whisker frequency. **a**, An Sp5l local population. **b,c**, Pooled data from all well localized local populations of cortical layer 4 barrels ($n = 9$) and layer 5a ($n = 6$) recorded during FM stimulations. In each station, pooled data is represented by the population response composed of all spikes generated by the different units recorded from the same station in different subjects and at different times. Spike times were coordinated according to onset times of the stimulus trains. Top,

PSTHs as a function of train time (vertical axis, PSTH time; horizontal, train time; colour axis, firing rate). Bottom, instantaneous stimulus frequencies (black), response latencies (blue) and spike counts (red) as a function of train time. Means of 36 trials $\pm$ s.e.m. Standard errors of the latency were computed; for each stimulus cycle, from the inter-trial variability of the latency to the first spike. Firing rates, spike counts and latencies were smoothed along train time by a convolution with a triangle of area 1 and base of ± 1 cycle.

of brainstem activity, which is transferred along the paralemniscal pathway in parallel with the activity conveyed along the lemniscal pathway. These results are consistent with the 1–2-ms delay between VPM and POM activation by brainstem electrical stimulations¹⁶. Of course, POM activity following the onset of activation could be affected by mixed lemniscal and paralemniscal cortical feedback. However, the similarity between the transient and steady-state response patterns within each pathway (Figs 1, 2) indicates that the sustained activation may also be pathway-specific, and that it is dominated, at least to a first approximation, by closed thalamocortical loops in each pathway.

A striking feature of the paralemniscal response pattern is that while the onset latency increased, the offset latency remained constant, which caused a reduction in the duration of the response and hence reduced spike counts. The latency increments cannot simply be attributed to VPM–POM interactions. These interactions are inhibitory¹⁷, which predicts a positive correlation between the response strength of the VPM neurons and the latencies of the POM neurons, whereas we observed the opposite (Fig. 2). However, an alternative explanation, based on closed-loop operation of the paralemniscal thalamocortical system, is consistent with our results. This explanation suggests that thalamocortical circuits implement phase-locked loops, a computational algorithm that is optimal for temporal decoding^{18,19}.

An implementation of a POM–cortex phase-locked loop that is consistent with the findings presented here is based on a simple closed-loop thalamocortical circuit²⁰ that includes cortical inhibition^{20,21}, cortical oscillatory mechanisms^{22–24} and thalamic gating²⁵. Briefly, this circuit establishes a negative feedback loop in which the latency of the cortical oscillations determines the latency and spike count of the thalamic neurons, and the thalamic spike count determines the subsequent latency of the cortical oscillations. Within such a circuit, thalamic and cortical latencies do not depend directly on the brainstem latency, but rather on the latency of the cortical oscillations. Consistent with the model and results presented here, cortical oscillating neurons display increasing latencies with increasing stimulus frequencies²⁴ (see Supplementary Information for a more detailed description of the model).

When properly tuned, such a simple thalamocortical circuit should be able to track changes in the input frequency with a delay of a single cycle, and to represent these changes in the latencies and spike counts of its thalamic and cortical neurons¹⁸ (see Supplementary Information). To achieve this, thalamic neurons must function in their gating mode. However, following a quiescent period, thalamic neurons shift to a non-gating mode in which their outputs do not appear to depend on the cortical signal²⁵. This can account for the relatively short latency of POM neurons to the first

stimulus cycle (Figs 1–3) and the relatively long period required for stabilization (Fig. 1). In the awake animal, thalamic neurons probably shift into a gating mode upon whisking¹³.

We tested the performance of the POM–cortex circuits in single-cycle temporal decoding by applying frequency-modulated (FM) stimuli as follows. To shift the thalamic neurons into a gating mode and to allow ‘locking in’ of the thalamocortical circuits¹⁸, the whisker stimulation frequency was first kept constant at 5 Hz for 2 s. Then the stimulus frequency was modulated for an additional 6 s (40% modulation depth, 0.4 Hz, initial phase 90°). The entire process was repeated 36 times for each simultaneous recording, and the responses of each local population were averaged across repetitions.

In the brainstem, the response to each stimulus cycle was the same, regardless of the instantaneous frequency (the reciprocal of the current period) of the stimulus (Fig. 4a). In contrast, once ‘locked in’ to the stimulus, paralemniscal thalamic and cortical neurons directly represented the input frequency, by both latency and spike-count, with almost no delay (see pooled layer 5a representation in Fig. 4c). The spike-count and latency representations were usually linearly related, reflecting the transformation of latency to spike count observed in this system. Linear regression estimates of spike counts from latencies could explain most of the spike count variability for layer 5a ($r^2 = 0.66 \pm 0.2$, $n = 6$, $P < 0.0001$) and POM ($r^2 = 0.51 \pm 0.12$, $n = 9$, $P < 0.0001$) local populations. As expected from the steady-state data, whisker frequency was represented by spike count alone in the lemniscal pathway (see pooled barrel representation in Fig. 4b). The lemniscal spike-count representation was not sensitive to the fast frequency transition (at $t = 2$ s) and exhibited a smaller dynamic range than the paralemniscal one (modulation depths of the spike-count representations were $12.5 \pm 15\%$ ($n = 13$) for VPM and layer 4 barrel neurons, and $25.5 \pm 14.5\%$ ($n = 15$) for POM and layer 5a neurons ($P < 0.05$, two-tailed t -test)). Thus, the temporal information of the stimulus was not reliably re-coded by lemniscal firing rates, even though the temporal information was faithfully replicated by lemniscal firing times.

During exploration and active touch, rats move their whiskers back and forth at frequencies between 4 and 11 Hz^{9,12,13,19}. These whisking movements yield spatial and temporal encoding of the external spatial information. For example, object location (relative to the face and initial position of whiskers) is encoded by the identity of the activated whiskers (spatial encoding) and by the temporal interval between protraction onset and perturbation by the object (temporal encoding)²⁶. Our results indicate that these two components may be processed separately and in parallel, in the lemniscal and paralemniscal pathways, respectively. The loss of time

locking to stimulus onset in the paralemniscal system is a reflection of the temporal-to-rate code translation process, in which latency shifts produce a rate-coded representation of the whisker frequency. The fixed time locking preserved along the lemniscal system probably reflects a lack of temporal processing within the whisking frequency range. On the other hand, the fixed time locking is required for reliable spatial processing during whisking, because, owing to whisker movements, temporal uncertainties in afferent firings result in uncertainties about whisker location. Thus, counterintuitively, fixed time locking is crucial for spatial processing, whereas variable latencies serve temporal processing. The response durations of the paralemniscal neurons are well tuned for temporal processing in the frequency range of whisking, as these response durations match the duration of a single protraction phase (> 50 ms) and thus allow the decoding of all possible temporal intervals during protraction. Note that decoding of temporal information at higher frequencies, such as those that encode the texture of objects scanned by the whiskers⁹, probably cannot be achieved by the slow paralemniscal system, but can be achieved by lemniscal thalamocortical circuits in which high frequencies can be coded (and translated to rate) by fine modulations of the response latency.

Our data indicate that the paralemniscal pathway of the rat somatosensory system is optimally tuned for temporal processing of vibrissal information around the whisking frequency range (~ 8 Hz), and that temporally encoded information can be translated to a rate code by thalamocortical loops. These findings point towards a parallel processing scheme in which spatially encoded (and perhaps high-frequency, temporally encoded) information is decoded by the lemniscal system, whereas low-frequency, temporally encoded information is decoded by the paralemniscal system. Sensory pathways dedicated to temporal processing have been reported in the electric fish and in the auditory brainstem of birds and mammals²⁷. Here we provide evidence for a separate sensory pathway for temporal processing that involves the thalamus and cortex. □

Methods

Animal procedures and electrophysiology

Experimental procedures were similar to those used previously^{24,28}. Briefly, 43 adult male Wistar albino rats (300 ± 25 g) obtained from the Animal Breeding Unit of The Weizmann Institute of Science were anaesthetized (urethane, 1.5 g kg^{-1} , injected intraperitoneally) and mounted in a modified stereotaxic device that allows free access to the somatosensory cortex and vibrissae. We monitored the depth of anaesthesia by assessing corneal and limb-withdrawal reflexes, and maintained it at stage III/3–4 (as determined by the dominant frequencies in the power spectra of local field potentials in the cortex¹⁴) by injections of urethane (10% of initial dose). Atropine methyl nitrate (0.3 mg kg^{-1} , intramuscular) was administered before general anaesthesia to prevent respiratory complications. Body temperature was maintained at $\sim 37^\circ\text{C}$.

The skull was exposed, and openings were made to allow electrode penetrations to different neuronal stations in the barrel cortex, thalamus and/or brainstem. Four to eight microelectrodes were driven into one or two stations simultaneously using one or two microdrive systems²⁸. We isolated single units by spike templates and multiple units by amplitudes using spike sorters (MSD-2; Alpha-Omega). We recorded 96 single-units and 62 multi-units in the brainstem, 333 and 234, respectively, in the thalamus, and 281 and 244, respectively, in the cortex. The number of neurons recorded by a single electrode (a local population) was estimated to be 5 ± 1.3 (mean $\pm$ s.d.) neurons. The care and treatment of rats were in accordance with the animal welfare guidelines of The Weizmann Institute of Science.

Whisker stimulation

Whisker stimulations consisted of pulses of compressed air, generated by a pneumatic pressure pump (Medical Systems) and delivered through 3 m of stiff, thick-walled tubing (6.5 mm outer diameter, 3.5 mm inner diameter). A micropipettor tip was attached to the end of the tubing to reduce its opening to a diameter of 0.7 mm and was positioned 10–20 mm caudally to the most caudal whiskers of the stimulated rows. We usually stimulated one or two rows of whiskers (including at least the four most caudal whiskers and associated straddlers). Constant-frequency stimuli were applied in blocks of 12×2 or 24 trains of 3 or 4 s each, with inter-train intervals of 2 or 1 s, respectively. FM stimuli were applied in blocks of 36 trains of 8 s each, with inter-train intervals of 2 s. The stimuli were delivered as constant pulses (see bottom trace in Fig. 1): during the first 50 ms of each cycle the whiskers were protracted (rise time, 30 ms; maximum pressure level at pump output,

0.7 kg cm^{-2}) and then allowed to retract (fall time, 10 ms). Pulse profiles were equal for all the frequencies used. Air-puff delays (from pump command to whisker movement) were calibrated using a microphone. In all experiments, latencies evoked by air puffs were compared with those evoked by a mechanical stimulator²⁸, which was attached to a single whisker 5–10 mm from the skin (amplitude, 560 μm ; rise and fall times, 5 ms). When mismatches (all were < 3 ms) were detected, the mechanical latencies were assumed.

Histology

We analysed the brains of the rats histologically to localize all recording sites. At the end of each recording session, we induced electrolytic lesions by passing currents ($3\text{--}5 \mu\text{A}$, 2×2 s, unipolar) through the tips of the electrodes. The brains were sectioned (coronally for cortical and thalamic recordings and parasagittally for brainstem recordings) and stained for cytochrome oxidase activity²⁸. In these preparations, the different cortical layers, barrel borders, thalamic and brainstem nuclei, lesions and electrode tracks were clearly visible (Fig. 1).

Data analysis

Only data from well localized recording sites (see Fig. 1) are described. We computed PSTHs using 1-ms bins and smoothed them by convolution with a triangle of area 1 and a base of ± 8 ms. Neuronal latencies were estimated from the smoothed PSTHs as the time in which the firing rate crossed the half-peak value (after subtracting background activity). Other latency estimations, such as latency to peak or to 0.1 peak value, yielded basically the same results. Spike counts were the average sum (per stimulus cycle) of spikes during the first 100 ms after stimulus onset.

Received 28 December 1999; accepted 1 May 2000.

- Woolsey, T. A. in *Encyclopedia of Neuroscience* (eds Adelman, G. & Smith, B.) 195–199 (Elsevier, Amsterdam, 1997).
- Diamond, M. E. & Armstrong-James, M. Role of parallel sensory pathways and cortical columns in learning. *Concepts Neurosci.* **3**, 55–78 (1992).
- Diamond, M. E., Armstrong-James, M. & Ebner, F. F. Somatic sensory responses in the rostral sector of the posterior group (POm) and in the ventral posterior medial nucleus (VPM) of the rat thalamus. *J. Comp. Neurol.* **318**, 462–476 (1992).
- Diamond, M. E., Armstrong-James, M., Budway, M. J. & Ebner, F. F. Somatic sensory responses in the rostral sector of the posterior group (POm) and in the ventral posterior medial nucleus (VPM) of the rat thalamus: dependence on the barrel field cortex. *J. Comp. Neurol.* **319**, 66–84 (1992).
- Hoogland, P. V., Welker, E., Van der Loos, H. & Wouterlood, F. G. in *Cellular Thalamic Mechanisms* (eds Bentivoglio, M. & Spreafico, R.) 152–162 (Elsevier, Amsterdam, 1988).
- Chmielowska, J., Carvell, G. E. & Simons, D. J. Spatial organization of thalamocortical and corticothalamic projection systems in the rat Sml barrel cortex. *J. Comp. Neurol.* **285**, 325–338 (1989).
- Lu, S. M. & Lin, R. C. Thalamic afferents of the rat barrel cortex: a light- and electron-microscopic study using *Phaseolus vulgaris* leucoagglutinin as an anterograde tracer. *Somatosens. Mot. Res.* **10**, 1–16 (1993).
- Koralek, K. A., Jensen, K. F. & Killackey, H. P. Evidence for two complementary patterns of thalamic input to the rat somatosensory cortex. *Brain Res.* **463**, 346–351 (1988).
- Carvell, G. E. & Simons, D. J. Biometric analyses of vibrissal tactile discrimination in the rat. *J. Neurosci.* **10**, 2638–2648 (1990).
- Hartings, J. A. & Simons, D. J. Thalamic relay of afferent responses to 1- to 12-Hz whisker stimulation in the rat. *J. Neurophysiol.* **80**, 1016–1019 (1998).
- Nicolelis, M. A. & Chapin, J. K. Spatiotemporal structure of somatosensory responses of many-neuron ensembles in the rat ventral posterior medial nucleus of the thalamus. *J. Neurosci.* **14**, 3511–3532 (1994).
- Welker, W. I. Analysis of sniffing of the albino rat. *Behaviour* **22**, 223–244 (1964).
- Fanselow, E. E. & Nicolelis, M. A. L. Behavioral modulation of tactile responses in the rat somatosensory system. *J. Neurosci.* **19**, 7603–7616 (1999).
- Friedberg, M. H., Lee, S. M. & Ebner, F. F. Modulation of receptive field properties of thalamic somatosensory neurons by the depth of anaesthesia. *J. Neurophysiol.* **81**, 2243–2252 (1999).
- Simons, D. J., Carvell, G. E., Hershey, A. E. & Bryant, D. P. Responses of barrel cortex neurons in awake rats and effects of urethane anaesthesia. *Exp. Brain Res.* **91**, 259–272 (1992).
- Chiaia, N. L., Rhoades, R. W., Fish, S. E. & Killackey, H. P. Thalamic processing of vibrissal information in the rat: II. Morphological and functional properties of medial ventral posterior nucleus and posterior nucleus neurons. *J. Comp. Neurol.* **314**, 217–236 (1991).
- Crabtree, J. W., Collingridge, G. L. & Isaac, J. T. A new intrathalamic pathway linking modality-related nuclei in the dorsal thalamus. *Nature Neurosci.* **1**, 389–394 (1998).
- Ahissar, E. Temporal-code to rate-code conversion by neuronal phase-locked loops. *Neural Comput.* **10**, 597–650 (1998).
- Kleinfeld, D., Berg, R. W. & O'Connor, S. M. Anatomical loops and their electrical dynamics in relation to whisking by rat. *Somatosens. Mot. Res.* **16**, 69–88 (1999).
- White, E. L. & Keller, A. Intrinsic circuitry involving the local axon collaterals of corticothalamic projection cells in mouse Sml cortex. *J. Comp. Neurol.* **262**, 13–26 (1987).
- Swadlow, H. A. Influence of VPM afferents on putative inhibitory interneurons in s1 of the awake rabbit—evidence from cross-correlation, microstimulation, and latencies to peripheral sensory stimulation. *J. Neurophysiol.* **73**, 1584–1599 (1995).
- Silva, L. R., Amitai, Y. & Connors, B. W. Intrinsic oscillations of neocortex generated by layer 5 pyramidal neurons. *Science* **251**, 432–435 (1991).
- Nicolelis, M. A. L., Baccala, L. A., Lin, R. C. S. & Chapin, J. K. Sensorimotor encoding by synchronous neural ensemble activity at multiple levels of the somatosensory system. *Science* **268**, 1353–1358 (1995).
- Ahissar, E., Haidarliu, S. & Zacksenhouse, M. Decoding temporally encoded sensory input by cortical oscillations and thalamic phase comparators. *Proc. Natl Acad. Sci. USA* **94**, 11633–11638 (1997).
- Sherman, S. M. & Guillery, R. W. Functional organization of thalamocortical relays. *J. Neurophysiol.* **76**, 1367–1395 (1996).

26. Zucker, E. & Welker, W. I. Coding of somatic sensory input by vibrissae neurons in the rat's trigeminal ganglion. *Brain Res.* **12**, 138–156 (1969).
27. Carr, C. E. Processing of temporal information in the brain. *Annu. Rev. Neurosci.* **16**, 223–243 (1993).
28. Haidarliu, S., Sosnik, R. & Ahissar, E. Simultaneous multi-site recordings and iontophoretic drug and dye applications along the trigeminal system of anesthetized rats. *J. Neurosci. Methods* **94**, 27–40 (1999).
29. Williams, M. N., Zahm, D. S. & Jacquin, M. F. Differential foci and synaptic organization of the principal and spinal trigeminal projections to the thalamus in the rat. *Eur. J. Neurosci.* **6**, 429–453 (1994).
30. Bourassa, J., Pinault, D. & Deschenes, M. Corticothalamic projections from the cortical barrel field to the somatosensory thalamus in rats: a single-fibre study using biocytin as an anterograde tracer. *Eur. J. Neurosci.* **7**, 19–30 (1995).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank M. Ahissar, and S. Barash for suggestions; A. Arieli, F. F. Ebner, K. O. Johnson, R. Malach and D. J. Simons for comments and discussions; M. E. Diamond for assistance during our initial thalamic recordings; and B. Schick for reviewing the manuscript. This work was supported by the US–Israel Binational Science Foundation and by the MINERVA Foundation, Germany. S.H. was supported by The Centre for Absorption of Scientists, Ministry of Absorption, Israel.

Correspondence and requests for materials should be addressed to Ehud Ahissar (e-mail: Ehud.Ahissar@weizmann.ac.il).

Presenilin is required for proper morphology and function of neurons in *C. elegans*

Nicole Wittenburg*, Stefan Eimer*, Bernard Lakowski*, Sascha Röhrig*, Claudia Rudolph† & Ralf Baumeister*

* Genzentrum, Ludwig-Maximilians-Universitaet, Feodor-Lynen-Strasse 25, D-81377 Munich, Germany

† EleGene GmbH, Am Klopferstspitz 19, D-82152 Martinsried, Germany

Mutations in the human presenilin genes cause the most frequent and aggressive forms of familial Alzheimer's disease (FAD)¹. Here we show that in addition to its role in cell fate decisions in non-neuronal tissues^{2–4}, presenilin activity is required in terminally differentiated neurons *in vivo*. Mutations in the *Caenorhabditis elegans* presenilin genes *sel-12* and *hop-1* result in a defect in the temperature memory of the animals. This defect is caused by the loss of presenilin function in two cholinergic interneurons that display neurite morphology defects in presenilin mutants. The morphology and function of the affected neurons in *sel-12* mutant animals can be restored by expressing *sel-12* only in these cells. The wild-type human presenilin PS1, but not the FAD mutant PS1 A246E, can also rescue these morphological defects. As *lin-12* mutant animals display similar morphological and functional defects to presenilin mutants, we suggest that presenilins mediate their activity in postmitotic neurons by facilitating Notch signalling. These data indicate cell-autonomous and evolutionarily conserved control of neural morphology and function by presenilins.

To study the activity of presenilin genes in neurons, we focused on *C. elegans sel-12* as the detailed morphology and function of many neurons in *C. elegans* is known. Like presenilins in other species, *C. elegans sel-12* is strongly expressed in neurons. We tested the functional integrity of the nervous system of *sel-12* mutants by looking for defects in the execution of a variety of behaviours, such as movement and response to mechanical, chemical and thermal stimuli.

We found that *sel-12* mutants display a highly penetrant defect in

their ability to sense and/or memorize temperature. Wild-type *C. elegans* display strong preference for their growth temperature, and can memorize it and store the information for several hours, suggesting a neuronal plasticity⁵. This behaviour can be studied with a simple experimental model. When placed in a radial thermal gradient on the agar surface of a petri dish, wild-type animals migrate to their preferred temperature, and then move in isothermal circles (Table 1, Fig. 1a). In contrast, the *sel-12(ar131)* and *sel-12(ar131)* and *sel-12(ar171)* mutant animals have lost the ability to perform isothermal tracks. Most animals are non-responsive to the temperature gradient and moved randomly on the plate (athermotactic behaviour), and 10% of the remaining animals moved to colder temperatures than the wild-type (cryophilic behaviour). These results indicate that *sel-12* mutants may have defects in the neural circuit for thermotaxis.

The neurons necessary for thermotaxis have been studied extensively by mutational analyses and laser ablation studies⁶. Temperature input activates the two AFD sensory neurons, which synapse extensively onto the two AIY interneurons. Chemical signals from AIY and AIZ (synaptic partners that represent the four central integrating interneurons), in turn, regulate postsynaptic inter- and motor neurons that control the motor response. We carefully examined the morphology of the AFD, AIZ and AIY neurons in *sel-12* animals using green fluorescent protein (GFP) reporter constructs, and saw no obvious defects in AFD and AIZ neurons (data not shown). However, we identified defects in the morphology of AIY neurons (Table 2, Fig. 2). In wild-type animals, the processes of both AIY neurons extend anteriorly from the cell bodies along the ventral cord, run around the nerve ring and meet and terminate at the dorsal midline⁷ (Fig. 2e). In adult *sel-12* mutants the AIY cell bodies are correctly positioned in the head ganglion. However, the AIY axons often grow too far anteriorly before turning and fasciculating in the nerve ring (Fig. 2d), and/or do not stop growth at the dorsal side of the nerve ring, but turn posteriorly, sometimes extending up to the midbody region (Fig. 2b–d, classified as severe defects in Table 2). In addition, short extra neurites often emerge directly from the cell soma or branch off the primary process (Fig. 2a, classified as minor defects in Table 2). The number of animals showing these types of defects in AIY morphology was higher in *sel-12(ar171)* mutants (35% defects) than in *sel-12(ar131)* mutants (20% defects; Table 2). This is consistent with the severity of these mutations and their effect on egg-laying behaviour (*ar171* carries a nonsense mutation in *sel-12* and genetically represents a null allele, and *ar131* carries a missense mutation in *sel-12*)². Behavioural defects are far more penetrant than the morphological defects, indicating that *sel-12* animals may also have more subtle defects in the AIY neurons than can be visualized with GFP

Table 1 Rescue of the *sel-12* thermotaxis defect

Genotype	Transgene	Fraction showing isothermal tracks
Wild type*		41/46
<i>sel-12(ar131)</i>		1/42
<i>sel-12(ar171)</i>		3/38
<i>hop-1(lg1501)</i>		0/44
<i>lin-12(n941)</i>		0/15
<i>hop-1(lg1501);sel-12(ar131)</i>		0/31
<i>hop-1(lg1501);sel-12(ar171)</i>		0/33
<i>sel-12(ar171);bys101</i>	<i>sel-12::sel-12</i>	38/49
<i>sel-12(ar171);bys100</i>	<i>sel-12::sel-12</i>	19/41
Wild type†		49/55
<i>sel-12(ar131)†</i>		3/40
<i>sel-12(ar171)†</i>		2/39
<i>sel-12(ar131); byEx103†</i>	<i>ttx-3::sel-12</i>	18/42
<i>sel-12(ar131); byEx115†</i>	<i>ttx-3::sel-12</i>	21/41
<i>sel-12(ar131); byEx115†</i>	<i>ttx-3::sel-12</i>	35/67
<i>sel-12(ar171); byEx115†</i>	<i>ttx-3::sel-12</i>	18/41

* Strain carries a *daf-6(e1377)* mutation that did not affect thermotaxis behaviour.

† Animals expressing *ttx-3::GFP*.

bys100 and *bys101* are independent chromosomally integrated arrays expressing *sel-12* cDNA from the *sel-12* promoter.

constructs. This is not unprecedented, because mutations in other axonal guidance genes often lead to highly penetrant behavioural defects with a much lower penetrance of morphological defects than is visible by light microscopy⁸.

To confirm that the defects we observed are due to a loss of *sel-12* activity, we transformed *sel-12* mutants with a *sel-12* complementary DNA under the control of the *sel-12* promoter. This construct rescued the egg-laying defect (data not shown), the thermotaxis behaviour of *sel-12(ar171)* and the neurite morphology defect (Tables 1 and 2). To determine whether *sel-12* activity is required cell-autonomously or non-cell-autonomously, we used the *ttx-3* promoter to express *sel-12* cDNA exclusively in AIY⁹. We analysed *sel-12(ar131)* and *sel-12(ar171)* mutants that expressed *ttx-3::sel-12* from an extra-chromosomal array. All transgenic animals still showed the fully penetrant egg-laying defect typical of *sel-12* mutants, but expression of *sel-12* solely in AIY restored isothermal tracking (Table 1, Fig. 1d). In addition, the morphology of the AIY neurons was indistinguishable from wild-type (Table 2). Together, our data indicate that SEL-12 activity in AIY is required cell-autonomously for correct neurite connectivity and for the only known function of this neuron.

A second somatically expressed presenilin, *hop-1*, exists in *C. elegans* and is thought to be largely redundant with *sel-12*. *hop-1* mutants were described to have no obvious phenotype on their own, but strongly enhance the developmental defect of *sel-12* mutants¹⁰; however, we find that *hop-1(lg1501)* mutants are also defective in their thermotaxis behaviour. None of the mutants that we tested performed isothermal tracking. Twenty-seven of the forty-four animals tested showed cryophilic behaviour similar to those animals in which both AIY neurons were ablated by microsurgery⁶, indicating that *hop-1* activity is required for AIY function (Fig. 1c). Like *sel-12* animals, *hop-1* animals displayed a weakly penetrant morphological defect in AIY (Table 2). We found that 91% of the *hop-1(lg1501);sel-12(ar131)* and 94% of the *hop-1(lg1501);sel-12(ar171)* double mutants showed abnormal AIY morphology. The severity of the defects was considerably stronger than that observed in the *hop-1(lg1501)*, *sel-12(ar131)* or *sel-12(ar171)* single mutants, indicating that the low penetrance of neuronal defects in *sel-12* mutants is due in part to the functional redundancy of *sel-12* and *hop-1*. These data also indicate that the activities of both presenilins are required for the function of the AIY interneurons and for their correct neurite morphology.

The vertebrate presenilin genes are strongly expressed in neurons and the protein has been detected in neurite growth cones and vesicles^{11–13}. This indicates that human presenilins may have a function in neurite outgrowth. To test whether human presenilins can functionally replace *sel-12* in the AIY neurons, we expressed the cDNAs of PS1 and the FAD mutant PS1 A246E under the control of a *sel-12* promoter in *sel-12* mutants. The neurite morphology in the strains expressing PS1 was completely rescued (Table 2). In contrast, rescue was reduced when we expressed PS1 A246E in *sel-12* mutants (Table 2). This is consistent with previous experiments showing that wild-type human presenilins, but not FAD mutants, can rescue the *sel-12* egg-laying defect^{3,4}. Our data indicate that human and *C. elegans* presenilin may have conserved functions in postmitotic neurons. We propose that in vertebrates, presenilins may also be required cell-autonomously to control neurite morphology.

As *sel-12* is expressed in most neurons in the *C. elegans* nervous system, we analysed the general neural anatomy of *sel-12* mutants using other GFP reporter genes, including a pan-neuronally expressed *unc-119::GFP* construct. The position and process morphology of most neurons appeared normal. However, irregular branching of neurites was sometimes observed in the BDU cells (a pair of interneurons of unknown function) and fasciculation and pathfinding defects were observed in neuronal processes in the ventral cord in 18% of the examined animals ($n = 103$). We also examined locomotion and tested many other behaviours (such as

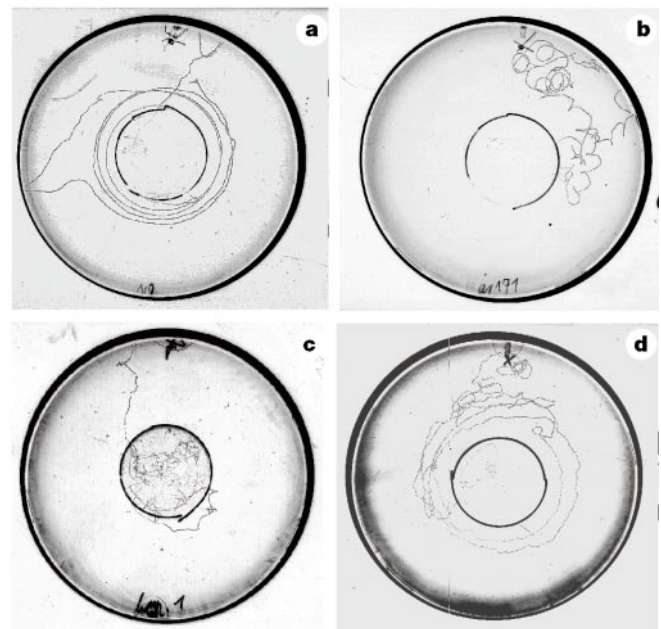


Figure 1 Tracking of animals on a radial temperature gradient. Animals raised at 20 °C were tested on a radial temperature gradient ranging from 17 °C (in the central circle) to room temperature at the edge. After the animals were transferred to the gradient (at the X-mark on top) they were given at least 1 hour to move freely on the plate. The tracks they left in the agar were photographed. A representative plate is shown for wild-type (a); *sel-12(ar171)* (b); *hop-1(lg1501)* animals (c); *sel-12(ar171)* animals expressing a *ttx-3::sel-12* transgene in AIY (d).

chemotaxis to food, foraging and mechanoperception of *sel-12* mutant animals) (data not shown), but could not identify any abnormalities in these behaviours. Accordingly, no anatomical defects were seen in the associated neurons. This indicates that, although *sel-12* is expressed in most or all neurons, it is absolutely required in only a subset of neurons.

Is the abnormal morphology of AIY a secondary consequence of abnormal AIY function? There is evidence that axon connectivity in *C. elegans* may not in all cases be determined by developmental outgrowth, but also can be refined by sensory activity¹⁴. To test whether AIY morphology may be affected by the lack of synaptic activity, we tested axon development in an *egl-19(n2368)* mutant that shows reduced neural transmission owing to a defective voltage-gated Ca^{2+} -channel subunit^{14,15}. Whereas only 1 of 137 newly hatched L1 animals tested displayed a non-wild-type AIY morphology, 38% of the *egl-19* animals showed AIY defects as adults. This indicates that lack of synaptic activity may influence AIY morphology. We compared AIY morphology in *sel-12(ar131)* and *sel-12(ar171)* mutants shortly after hatching (L1) with that of adult animals. Sixteen percent of L1 ($n = 167$) and 20% of adult *ar131* mutants, and 26% of L1 ($n = 173$) and 35% of adult *ar171* mutants exhibited AIY defects. As the defects can be observed shortly after hatching, and therefore before a significant and biologically relevant sensory input, we suggest that the morphological defects of *sel-12* mutants are predominantly generated during axonal outgrowth in AIY and not as a consequence of a reduced activity of their neurons.

Results in a variety of model systems have shown that presenilin activity controls the function of several, obviously non-related transmembrane proteins. Presenilin activity is required for the proteolytic processing of the amyloid precursor protein involved in Alzheimer's disease¹, but also for the proteolysis of the Irf1 receptor¹⁶ and Notch receptor^{17–19}. In *C. elegans*, presenilins are required for Notch-mediated cell fate decisions in the vulval and uterine tissues². The *C. elegans* Notch homologues LIN-12 and GLP-1 affect neural development by controlling lineage identity, but

no expression or activity of *lin-12* or *glp-1* in the neurons involved in the circuitry required for thermotaxis has been reported.

To determine whether the consequences of reduced Notch activity in AIY are similar to those of mutations in *sel-12*, we examined the neural anatomy of AIY in *lin-12* and *glp-1* strong loss-of-function alleles. We observed a very weak AIY defect in *glp-1* alleles. However, there was a highly penetrant defective AIY morphology in *lin-12(n941)* mutants (61% defects, Table 1). The same morphological abnormalities were found as in *sel-12* mutant animals, including a failure of the AIY axon to terminate at the dorsal side of the nerve ring and a remarkable increase in neurite sprouting very similar to that seen in *sel-12* mutants (Fig. 2f). As with *sel-12* mutants, *lin-12* animals also display a thermotaxis defect. Although most of the *lin-12(n941)* animals displayed strongly impaired movement on a temperature gradient, 15 out of 44 animals did migrate. All of these animals moved to the centre of the radial temperature gradient and may be considered cryophilic (Fig. 1c). From these data we conclude that the reduction of LIN-12/Notch activity also results in a morphological and behavioural defect of AIY. The similarity of the observed phenotype indicates that presenilins may control AIY morphology and function through the activity of LIN-12/Notch. The defective motility of *lin-12(n941)* may also indicate additional neurogenic or myogenic roles for *lin-12* that have not been analysed in detail.

We have shown that the AIY defects in *sel-12* mutants can be fully rescued by human PS1, indicating that the neural function of presenilins is evolutionarily conserved. As we have suggested for *C. elegans*, presenilins in other organisms may control the connectivity and function of selected neurons by modulating Notch signalling in these cells. It has been shown that the *Drosophila* Notch is involved in specific regionally constricted guidance of neuronal outgrowth²⁰, although the cells that require Notch activity were not determined. Experiments in cell culture show that Notch signalling affects neurite growth^{21,22}, and that a functional interaction between human PS1 and Notch1 exists in postmitotic mammalian neurons²³. On the basis of cell-culture experiments, both autonomous and non-autonomous regulation of neurite development by Notch signalling have been suggested²⁴. Our results indicate a conserved function for presenilins in postmitotic neurons which is supported by data from cell-culture experiments. Owing to the lack of obvious other behavioural defects of *sel-12* mutants, it is likely that *sel-12* activity is only required in a subset of the neurons in *C. elegans*, despite the broad expression of *sel-12* in the nervous system. It is interesting to note that the AIY neurons produce the neurotransmitter acetylcholine (J. Duerr and J. Rand, personal communication) and that cholinergic neurons have an increased susceptibility to dysfunction in Alzheimer's disease²⁵. Mutations in the *C. elegans* choline acetyl transferase *cha-1* also affect AIY neurite

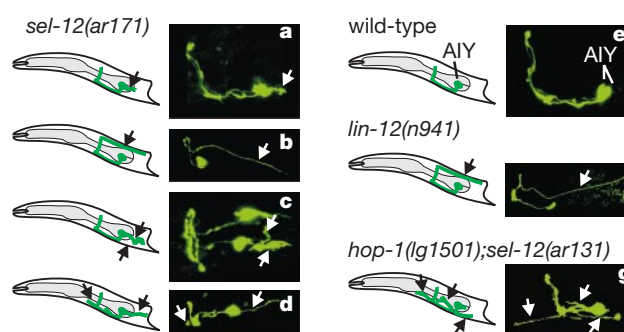


Figure 2 AIY interneuron morphology in *sel-12*, wild-type, *lin-12* and *hop-1; sel-12* mutant animals. Left, drawing of morphology of AIY interneurons (adapted from ref. 7). Right, expression of *tx-3::GFP* in AIY. Abnormal axonal projections are indicated by arrows. **a–d**, Defects seen in *sel-12* mutants. **a**, A posteriorly directed process sprouts from the cell soma. **b**, An axon does not terminate at dorsal mid-line but continues to grow posteriorly; in severe cases we saw extension to the midbody region. **c**, Rootlet-like projections extend from the cell somata in no defined direction. **d**, An axon leaves the fascicle when it turns into the nerve ring. Additional rootlet-like neurites project posteriorly from the cell soma. **e**, Wild-type morphology. AIY processes run anteriorly up the ventral cord, then enter the nerve ring and finally meet and terminate at the dorsal mid-line. **f**, Defects observed in *lin-12(n941)* animals resemble those seen in *sel-12* mutants. **g**, In *hop-1(lg1501); sel-12(ar131)* AIY defects also resemble the defects seen in the other mutants, but were more severe. Small axonal projections were scored as minor defects (**a**). All defects that lead to a gross change in AIY morphology are grouped as severe defects (**b–d, f, g**).

morphology (our unpublished results), and it has been reported that mutations in the human presenilins can reduce choline acetyl transferase activity that may contribute to the cognitive impairment in Alzheimer's disease²⁵. In the human brain, the neural structure whose function is most severely affected by presenilin mutations during the progression of FAD is the hippocampus, which is central to the perception and long-term storage of memory²⁶. Remarkably, the neurons most strongly affected in *C. elegans* by presenilin mutations are the AIY neurons that are involved in the only neural circuit for which synaptic plasticity has been suggested⁶.

Here we correlate behavioural and morphological assays to monitor presenilin function in a pair of *C. elegans* neurons. The detailed knowledge of the *C. elegans* neural anatomy and the conservation of presenilin function can now be exploited to analyse the molecular details leading to these defects. Understanding the role of presenilins in the nervous system may result in an understanding of the abnormal neurite pattern and neural deficits of Alzheimer's disease patients. □

Table 2 AIY morphology defects

Strain*	Transgene	No defects (%)	Minor defects (%)	Severe defects (%)	No. of animals
Wild type		97	3	0	113
<i>sel-12(ar131)</i>		80	16	4	159
<i>sel-12(ar171)</i>		65	12	23	145
<i>hop-1(lg1501)</i>		70	17	13	100
<i>lin-12(n941)</i>		39	43	18	108
<i>glp-1(e2144)</i>		90	6	4	220
<i>glp-1(q46)</i>		92	7	1	112
<i>hop-1(lg1501); sel-12(ar131)</i>		4	27	69	70
<i>hop-1(lg1501); sel-12(ar171)</i>		9	65	26	43
<i>sel-12(ar171)</i>	<i>sel-12::sel-12</i>	96	4	9	102
<i>sel-12(ar131)</i>	<i>tx-3::sel-12</i>	98.0 ± 2.0	1.0 ± 1.0	1.0 ± 1.0	337
<i>sel-12(ar171)</i>	<i>tx-3::sel-12</i>	98.0 ± 2.0	1.0 ± 1.0	1.0 ± 1.0	342
<i>sel-12(ar131)</i>	<i>sel-12::PS1</i>	98.5 ± 0.5	0.5 ± 0.5	1.0 ± 0.0	369
<i>sel-12(ar171)</i>	<i>sel-12::PS1</i>	94.5 ± 2.5	3.5 ± 1.5	2.0 ± 1.0	336
<i>sel-12(ar131)</i>	<i>sel-12::PS1 A246E</i>	78.0 ± 6.0	14.0 ± 6.0	8.0 ± 0.0	224
<i>sel-12(ar171)</i>	<i>sel-12::PS1 A246E</i>	74.5 ± 4.5	15.5 ± 6.5	10.0 ± 2.0	453

The neuronal and axonal morphology of AIY in *C. elegans* wild-type, presenilin, Notch and *sel-12* transgenic animals was examined. The percentage of animals displaying no defects, minor defects and severe defects (for definition see Fig. 2) are listed along with the total number of animals examined. All values for the transgenic lines are given as the mean percentage of two independent lines ± the standard deviation.

* Strains carry an extrachromosomal array containing *tx-3::GFP*.

Methods

General procedures

C. elegans strains were grown as described²⁷. *sel-12(ar171)* was separated from the marker mutation *unc-1(e538)* by standard techniques to allow analysis of thermotaxis. *sel-12(ar131)* and *sel-12(ar171)* were each backcrossed five times. XY1002 *hop-1(lg1501)* contains a 1,878-base-pair (bp) deletion in the gene. It eliminates bp 20,359–22,238 of cosmid C18E3 leading to a deletion of the carboxy-terminal exons 4–6 of *hop-1*.

Behavioural assays

We performed single animal thermotaxis assays² and the results were scored⁶ as described.

Expression constructs and generation of transgenic animals

sel-12 cDNA was obtained by PCR from a *C. elegans* cDNA library (Genbank accession number AF171064). This cDNA was cloned under the control of a 2.8-kilobase (kb) fragment 5' of the translational start ATG of *sel-12* to generate pBY895. pBY478 contains a *KpnI*–*SpeI* fragment of the *ttx-3* promoter²⁹ that drives expression of the *sel-12* cDNA exclusively in the AIY neurons. The expression constructs for human PS1 (pBY140) and the FAD mutant PS1 A246E (pBY147) have been described⁴. pBY140, pBY147 and pBY478 were injected into *him-8(e1489)* with *ttx-3::GFP* as a transformation marker. Two independent lines per construct were crossed into *sel-12(ar131)* and *sel-12(ar171)*. pBY895 was directly injected into *sel-12(ar171)* using *tph-1::GFP* as a transformation marker. Two integrated lines were obtained after X-ray irradiation and subsequently outcrossed five times each (*byIs100* and *byIs101*).

Identification of neuroanatomy

The *unc-119::GFP* reporter construct is expressed in all neurons²⁸. To identify AFD, a cell-specific *gcy-8::GFP* was used²⁹. AIY morphology was observed by using *ttx-3::GFP*⁹. We generated seven independent transgenic lines by injecting different *ttx-3::GFP* concentrations (1–50 ng μL^{-1}). All lines displayed AIY morphology defects in a *sel-12(ar171)* background. Line *byEx116* was chosen for all subsequent analyses, and was crossed into the different genetic backgrounds. AIZ was identified by using *lin-11::GFP* as a marker³⁰. Additional GFP reporter genes used contain *mec-7*, *tph-1* and *sel-12* promoters, which drive GFP expression in different subsets of neurons. All GFP constructs were injected into wild-type animals and subsequently crossed into various mutant backgrounds. *lin-12(n941);byEx116* were obtained by crossing into MT1965 *lin-12(n941)/eT1 III*; *him-5(e1467)/eT1[him-5(e1476)]* V. Animals homozygous for *lin-12* were Pvl (protruding vulva phenotype 1) and sterile. Homozygous *glp-1(e2144)* *byEx116* animals were identified on the basis of their embryonic lethality at 25 °C. *byEx116* was also crossed into *hop-1(lg1501);sel-12(ar131)* and *hop-1(lg1501);sel-12(ar171)* strains. For this purpose, *hop-1(lg1501)* was balanced over *unc-73(e936)*. Double homozygous animals were identified as being sterile and Pvl (ref. 10). The neuronal and axonal morphology was studied at 630–1,000-fold magnification using a Zeiss Axioskop compound microscope. Abnormal axonal projections were documented using a Leica TCS NT confocal microscope.

Received 11 February; accepted 25 May 2000.

- Haass, C. & De Strooper, B. The presenilins in Alzheimer's disease—proteolysis holds the key. *Science* **286**, 916–919 (1999).
- Levitani, D. & Greenwald, I. Facilitation of *lin-12*-mediated signalling by *sel-12*, a *Caenorhabditis elegans* S182 Alzheimer's disease gene. *Nature* **377**, 351–354 (1995).
- Levitani, D. *et al.* Assessment of normal and mutant human presenilin function in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **93**, 14940–14944 (1996).
- Baumeister, R. *et al.* Human presenilin-1, but not familial Alzheimer's disease (FAD) mutants, facilitate *Caenorhabditis elegans* Notch signalling independently of proteolytic processing. *Genes Funct.* **1**, 149–159 (1997).
- Hedgecock, E. M. & Russell, R. L. Normal and mutant thermotaxis in the nematode *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **72**, 4061–4065 (1975).
- Mori, I. & Ohshima, Y. Neural regulation of thermotaxis in *Caenorhabditis elegans*. *Nature* **376**, 344–348 (1995).
- White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Phil. Trans. R. Soc. Lond.* **314**, 1–340 (1986).
- Hedgecock, E. M., Culotti, J. G., Thomson, J. N. & Perkins, L. A. Axonal guidance mutants of *Caenorhabditis elegans* identified by filling sensory neurons with fluorescent dyes. *Dev. Biol.* **111**, 158–170 (1985).
- Hobert, O. *et al.* Regulation of interneuron function in the *C. elegans* thermoregulatory pathway by the *ttx-3* LIM homeobox gene. *Neuron* **19**, 345–357 (1997).
- Westlund, B., Parry, D., Basson, M. & Johnson, C. D. Reverse genetic analysis of *Caenorhabditis elegans* presenilins reveals redundant but unequal roles for *sel-12* and *hop-1* in Notch-pathway signalling. *Proc. Natl Acad. Sci. USA* **96**, 2497–2502 (1999).
- Kovacs, D. M. *et al.* Alzheimer-associated presenilins 1 and 2: Neuronal expression in brain and localization to intracellular membranes in mammalian cells. *Nature Med.* **2**, 224–229 (1996).
- Cook, D. G. *et al.* Expression and analysis of presenilin 1 in a human neuronal system: Localization in cell bodies and dendrites. *Proc. Natl Acad. Sci. USA* **93**, 9223–9228 (1996).
- Lee, M. K. *et al.* Expression of Presenilin 1 and 2 (PS1 and PS2) in human and murine tissues. *J. Neurosci.* **16**, 7513–7525 (1996).
- Peckol, E., Zallen, J., Yarrow, J. & Bargmann, C. Sensory activity affects sensory axon development in *C. elegans*. *Development* **126**, 1891–1902 (1999).
- Lee, R., Lobel, L., Hengartner, M., Horvitz, H. & Avery, L. Mutations in the $\alpha 1$ subunit of an L-type voltage-activated Ca^{2+} channel cause myotonia in *Caenorhabditis elegans*. *EMBO J.* **16**, 6066–6076 (1997).

- Katayama, T. *et al.* Presenilin-1 mutations downregulate the signalling pathway of the unfolded-protein response. *Nature Cell Biol.* **1**, 479–485 (1999).
- Struhl, G. & Greenwald, I. Presenilin is required for activity and nuclear access of Notch in *Drosophila*. *Nature* **398**, 522–525 (1999).
- Wolfe, M. S. *et al.* Two transmembrane aspartates in presenilin-1 required for presenilin endo-proteolysis and gamma-secretase activity. *Nature* **398**, 513–517 (1999).
- Ye, Y., Lukinova, N. & Fortini, M. E. Neurogenic phenotypes and altered Notch processing in *Drosophila* presenilin mutants. *Nature* **398**, 525–529 (1999).
- Giniger, E., Jan, L. & Jan, Y. Specifying the path of the intersegmental nerve of the *Drosophila* embryo: a role for Delta and Notch. *Development* **117**, 431–440 (1993).
- Berezovska, O. *et al.* Notch1 inhibits neurite outgrowth in postmitotic primary neurons. *Neuroscience* **93**, 433–439 (1999).
- Sestan, N., Artavanis-Tsakonas, S. & Rakic, P. Contact-dependent inhibition of cortical neurite growth mediated by Notch signaling. *Science* **286**, 741–746 (1999).
- Berezovska, O. *et al.* The Alzheimer-related gene presenilin 1 facilitates notch 1 in primary mammalian neurons. *Brain Res. Mol. Brain Res.* **69**, 273–280 (1999).
- Franklin, J. L. *et al.* Autonomous and non-autonomous regulation of mammalian neurite development by Notch1 and Delta. *Curr. Biol.* **9**, 1448–1457 (1999).
- Pedersen, W. A., Guo, Q., Hartman, B. K. & Mattson, M. P. Nerve growth factor-independent reduction in choline acetyltransferase activity in PC12 cells expressing mutant presenilin-1. *J. Biol. Chem.* **272**, 22397–22400 (1997).
- Mayford, M. & Kandel, E. R. Genetic approaches to memory storage. *Trends Genet.* **15**, 463–470 (1999).
- Brenner, S. The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71–94 (1974).
- Maduro, M. & Pilgrim, D. Identification and cloning of *unc-119*, a gene expressed in the *Caenorhabditis elegans* nervous system. *Genetics* **141**, 977–988 (1995).
- Yu, S., Avery, L., Baude, E. & Garbers, D. Guanylyl cyclase expression in specific sensory neurons: a new family of chemosensory receptors. *Proc. Natl Acad. Sci. USA* **94**, 3384–3387 (1997).
- Hobert, O., D'Alberti, T., Liu, X. & Ruvkun, G. Control of neural development and function in a thermoregulatory network by the LIM homeobox gene *lin-11*. *J. Neurosci.* **18**, 2084–2096 (1998).

Acknowledgements

Some strains used in this study were provided by the *Caenorhabditis* Genetics Center, which is funded by the NIH National Center for Research Resources. We thank I. Greenwald, E. Lambie and H. R. Horvitz for strains; R. Barstead for cDNA libraries; O. Hobert and A. Fire for plasmids and R. Donhauser and M. Grim for technical assistance. We also thank I. Mori and H. Kagoshima for their help with interpreting thermotaxis data, and C. Haass and the members of our lab for discussions and for critically reading the manuscript. Part of this work was supported by grants from the DFG to R.B. and from EMBO to B.L.

Correspondence and requests for materials should be addressed to R.B. (e-mail: bmeister@lmb.uni-muenchen.de).

A chemokine-driven positive feedback loop organizes lymphoid follicles

K. Mark Ansel^{*†}, Vu N. Ngo^{*†}, Paul L. Hyman^{*}, Sanjiv A. Luther^{*}, Reinhold Förster[‡], Jonathon D. Sedgwick[§], Jeffrey L. Browning^{||}, Martin Lipp[‡] & Jason G. Cyster^{*}

^{*} Department of Microbiology and Immunology, University of California San Francisco, San Francisco, California 94143, USA

[‡] Molecular Tumorigenetics and Immunogenetics, Max-Delbrück-Center for Molecular Medicine, 13092 Berlin, Germany

[§] DNAX Research Institute, Palo Alto, California 94304, USA

^{||} Biogen Inc., Cambridge, Massachusetts 02142, USA

[†] These authors contributed equally to this work

Lymphoid follicles are B-cell-rich compartments of lymphoid organs that function as sites of B-cell antigen encounter and differentiation. CXC chemokine receptor-5 (CXCR5) is required for B-cell migration to splenic follicles¹, but the requirements for homing to B-cell areas in lymph nodes remain to be defined. Here we show that lymph nodes contain two types of B-cell-rich compartment: follicles containing follicular dendritic cells, and areas lacking such cells. Using gene-targeted mice, we establish that B-lymphocyte chemoattractant (BLC/BCA1)^{2,3} and its receptor,

CXCR5, are needed for B-cell homing to follicles in lymph nodes as well as in spleen. We also find that BLC is required for the development of most lymph nodes and Peyer's patches. In addition to mediating chemoattraction, BLC induces B cells to up-regulate membrane lymphotoxin $\alpha 1\beta 2$, a cytokine that promotes follicular dendritic cell development and BLC expression^{4,5}, establishing a positive feedback loop that is likely to be important in follicle development and homeostasis. In germinal centres the feedback loop is overridden, with B-cell lymphotoxin $\alpha 1\beta 2$ expression being induced by a mechanism independent of BLC.

The BLC gene was inactivated by deletion of a portion of exon 2 which encodes amino acids 27–60, including 3 of the 4 conserved cysteine residues (Fig. 1a). Mice homozygous for the targeted locus (Fig. 1b) lacked detectable BLC messenger RNA (Fig. 1c) and protein (Fig. 1d). Anatomical analysis revealed that BLC-deficient mice had severe but incompletely penetrant defects in development of peripheral lymphoid organs. Most mice lacked inguinal, iliac, sacral, brachial and axillary lymph nodes, among others (Table 1); however, most of these lymph nodes were found at varying low frequencies, and several other lymph nodes developed normally, with all animals possessing a full set of mesenteric lymph nodes (Table 1). The lymphoid patch of the caecum was also absent and the number of Peyer's patches was severely reduced (Table 1). Those Peyer's patches that were found were typically smaller and lacked the characteristic multi-domed structure of wild-type Peyer's patches. Previous characterization of CXCR5-deficient mice established that they typically lack inguinal lymph nodes and have a less severe deficiency in Peyer's patches than do BLC-deficient mice¹. Further analysis of animals on a 129 or 129 $\times$ B6 mixed background indicated that they often lack additional peripheral lymph nodes (Table 1). These findings establish a broad requirement for the BLC–CXCR5 ligand–receptor pair in one of the earliest steps in the development of Peyer's patches and most lymph nodes.

To test whether BLC is needed for organization of B cells in lymphoid follicles, sections of lymph nodes and spleen from BLC^{−/−} animals were analysed for B-cell distribution. In all these organs, B cells failed to organize in polarized follicular clusters, and instead appeared as a ring of cells at the perimeter of T-cell areas (Fig. 2a). The boundary between B-cell-rich areas and T zones was often poorly demarcated, with increased numbers of B cells and T cells in reciprocal areas (Fig. 2a, b). In addition, staining spleen sections for immunoglobulin (Ig)M and IgD revealed a thickened ring of IgM^{hi}IgD^{lo} marginal zone B cells (Fig. 2a, top). The segregation of B cells between this area and the inner ring of B cells was also disrupted, with increased numbers of IgM^{lo}IgD^{hi} B cells located in the outer marginal zone (Fig. 2a, top).

A key property of B-cell follicles is the presence of immune-complex-presenting follicular dendritic cells (FDCs)⁶. Staining for complement receptor-1 (CR1), which is highly expressed on FDCs and marginal zone B cells, revealed an absence of primary follicle FDCs in BLC-deficient spleen and lymph nodes (Fig. 2b). Previous studies in CXCR5-deficient mice established a requirement for this

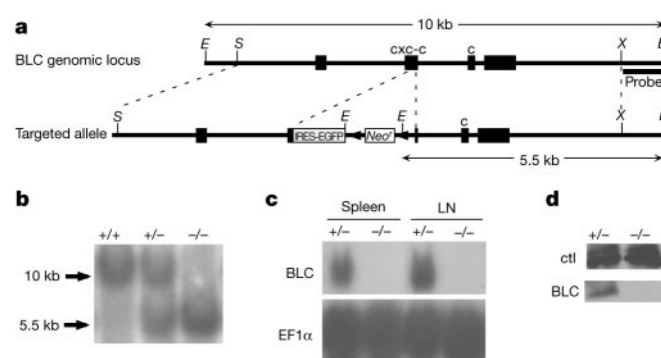


Figure 1 BLC genomic locus and lack of BLC mRNA and protein in BLC^{−/−} mice. **a**, BLC genomic locus before and after gene targeting. Black boxes indicate exons 1–4. The four cysteine residues are encoded by exons 2 and 3, as indicated. E, *Eco*R1; S, *Sac*I; X, *Xho*I. **b**, Southern blot analysis of *Eco*R1-digested genomic DNA from +/+, +/- and -/- mice. **c**, Northern blot of total RNA from +/- and -/- spleen and mesenteric lymph node. **d**, Western blot of heparin precipitates from +/- and -/- spleen lysates for BLC. ctl, non-specific band detected with anti-BLC serum, shown as a loading control.

receptor in B-cell follicle formation in spleen and Peyer's patches, but did not indicate that the receptor has a role in lymph nodes¹. During our analysis of BLC expression in lymph nodes, however, we observed two types of B-cell-rich zones: polarized follicles that are BLC positive and contain FDCs but few T cells; and areas that lack staining for BLC or FDC markers, contain substantial numbers of T cells, and lack a polarized morphology (Fig. 2c, see also e, f). BLC staining was detected on a subset of CR1⁺ FDCs and on adjacent CR1[−] follicular stromal cells (Fig. 2d). Analysis of lymph nodes from CXCR5^{−/−} animals for these two types of B-cell areas revealed that they lack FDC-containing primary follicles (data not shown). In transfer experiments, CXCR5^{+/+} B cells localized in FDC-containing follicles and FDC-deficient B-cell areas of wild-type lymph nodes (Fig. 2e). By contrast, CXCR5^{−/−} B cells failed to enter FDC-containing follicles (Fig. 2f). CXCR5^{−/−} B cells also did not migrate into follicles in the spleen and Peyer's patches, as previously observed¹; however, CXCR5^{−/−} B cells did localize in the FDC-deficient B-cell-rich areas in lymph nodes (Fig. 2f). B-cell homing to lymph node follicles therefore requires BLC and CXCR5, but neither molecule is necessary for migration to FDC-deficient B-cell-rich regions of secondary lymphoid tissues.

A lack of FDCs and organized follicles, as well as impaired development of several lymphoid organs, are phenotypes shared between BLC- and lymphotoxin (LT)-deficient mice⁴. Cell transfer experiments showed that B cells are an essential source of LT $\alpha 1\beta 2$ for the development of FDCs, but it was not clear whether LT $\alpha 1\beta 2$ was expressed by naive B cells^{7–9}. Although B cells express LT β constitutively^{10–12}, surface expression of this subunit requires LT α , and LT α expression has only been reported after exposure to activating stimuli, such as CD40L or interleukin (IL)-4 (ref. 11),

Table 1 Defective lymph node and Peyer's patch development in BLC and CXCR5-deficient mice

Genotype	Lymph nodes*												Peyer's patches†			
	Inguinal	Iliac/ periaortic	Axillary	Brachial	Popliteal	Deep cervical	Superficial cervical	Facial	Renal	Mesenteric	Sacral/caudal	Parathyroid	0	1–2	3–6	>6
BLC ^{+/+}	41(0)/41	38(0)/38	42(0)/42	42(0)/42	15(2)/19	14(2)/16	47(0)/47	47(0)/47	12(3)/15	50/50	37/38	15/15	0	0	2	29
BLC ^{−/−}	0(2)/41	0(0)/41	1(5)/42	1(8)/42	0(0)/18	1(2)/14	42(1)/46	41(2)/46	0(0)/12	50/50	1/40	0/14	21	11	5	0
CXCR5 ^{+/+}	0(0)/5	0(0)/5	2(0)/5	2(0)/5	0(0)/5	0(0)/5	3(1)/5	0(0)/5	5/5	5/5	0/5	0/5	1	2	1	1
CXCR5 ^{−/−} ‡	0(0)/10	0(0)/10	0(1)/10	0(1)/10	10(0)/10	0(0)/10	9(1)/10	10(0)/10	3(3)/9	10/10	1/10	0/10	0	5	3	1

BLC^{+/+}, BLC^{+/−}, BLC^{−/−} and CXCR5^{−/−} mice on a mixed B6/129 strain background were examined for the presence of lymph nodes and Peyer's patches.

* For lymph nodes, data shown are of the form *x*/*n*, where *x* is the number of mice in which the respective lymph nodes were identified out of *n* mice examined. For nodes that develop in two symmetrical sites, the number in parentheses represents the number of mice in which a lymph node had developed in one of the two sites. The complete absence of lymph nodes was confirmed in some mice by intraperitoneal injection of 300 μ l 1% Pontamine Blue dye in sterile PBS 10–20 days before dissection. The dye becomes concentrated in lymphoid organs as it is cleared from surrounding tissues, allowing easy detection of even minute lymphoid aggregates³⁰.

† For Peyer's patches, data shown are the number of mice in which the indicated number of PPs was identified along the small intestine.

‡ Ten additional CXCR5^{−/−} mice on the 129 strain background (six times backcrossed) were examined, revealing strain-dependent differences in lymph node development.

or treatment with endotoxin or phorbol esters¹³. Development of primary follicles is not associated with B-cell activation, however, as it occurs in germ-free mice¹⁴, Ig-transgenic mice¹⁵, T-cell-deficient mice¹⁶, CD40-deficient mice¹⁷ and IL4-deficient mice¹⁸. To investigate LT α 1 β 2 expression on naive B cells, we used a soluble form of the LT α 1 β 2 receptor, LT β R-Ig, to stain freshly isolated cells (Fig. 3a–g). Notably, a significant fraction of total B cells from wild-type spleen showed low but detectable LT β R-Ig binding compared with LT α - or LT β -deficient spleen cells (Fig. 3c, d, g). An even larger fraction of lymph node B cells stained with LT β R-Ig (Fig. 3e, f, g), whereas B cells in the blood (Fig. 3a, b, g) and bone marrow (data not shown) were mostly negative. Similar levels of LT α 1 β 2 were observed on B cells in Ig-transgenic animals (Fig. 3g), showing that the expression was not induced by antigen. We then tested LT α 1 β 2 expression on B cells in BLC-deficient animals (Fig. 3h). LT α 1 β 2 expression on B cells in BLC^{-/-} peripheral blood B cells was not different from wild-type controls, but expression on B cells from spleen and lymph nodes was significantly reduced (Fig. 3h), establishing that BLC is required for normal LT α 1 β 2 expression on naive B cells.

To determine whether recirculating B cells upregulated LT α 1 β 2 expression as they moved into BLC-expressing lymphoid organs, we isolated cells from the blood of Ly5.1⁺ donor mice and transferred them into congenic Ly5.2⁺ recipients. At 0.5 and 6 h after transfer, recipient splenocytes were collected and stained with anti-Ly5.1 to detect the transferred cells, and with LT β R-Ig (Fig. 3i). At the early time point, when most of the transferred B cells in the spleen are in

the non-lymphoid area (data not shown), few of the cells expressed LT α 1 β 2 (Fig. 3i). Six hours after transfer, however, when many of the transferred cells are in BLC⁺ follicular areas (data not shown), LT α 1 β 2 expression was readily detectable (Fig. 3i). When B cells were taken from spleen and transferred intravenously to syngeneic recipients, LT β R-Ig staining of cells in recipient blood and spleen soon after transfer had diminished to levels similar to endogenous blood B cells (Fig. 3i). The mechanism for this downregulation is not yet understood, but it suggests that as B cells leave a lymphoid organ and enter the blood they rapidly lose surface LT α 1 β 2 expression. In practical terms, this finding allowed us to use cells from spleen for further transfer experiments to test the mechanism of LT α 1 β 2 upregulation on cells entering lymphoid organs. Within 6 h of transfer, untreated spleen B cells expressed amounts of LT α 1 β 2 equal to endogenous B cells (Fig. 3i). By contrast, cells pretreated with pertussis toxin (PTX), an inhibitor of signalling by G α i-coupled receptors including all known chemokine receptors, failed to upregulate LT α 1 β 2 expression after transfer (Fig. 3i). Furthermore, minimal LT α 1 β 2 upregulation occurred when B cells from wild-type mice were transferred to BLC-deficient recipients (Fig. 3j, k). Notably, when B cells from BLC-deficient donors were transferred to wild-type recipients an exaggerated induction of LT α 1 β 2 occurred (Fig. 3l, m), perhaps because B cells that develop in BLC-deficient animals are hypersensitive to the chemokine. Together, these findings provide strong support for the hypothesis that recirculating B cells upregulate LT α 1 β 2 as they migrate to

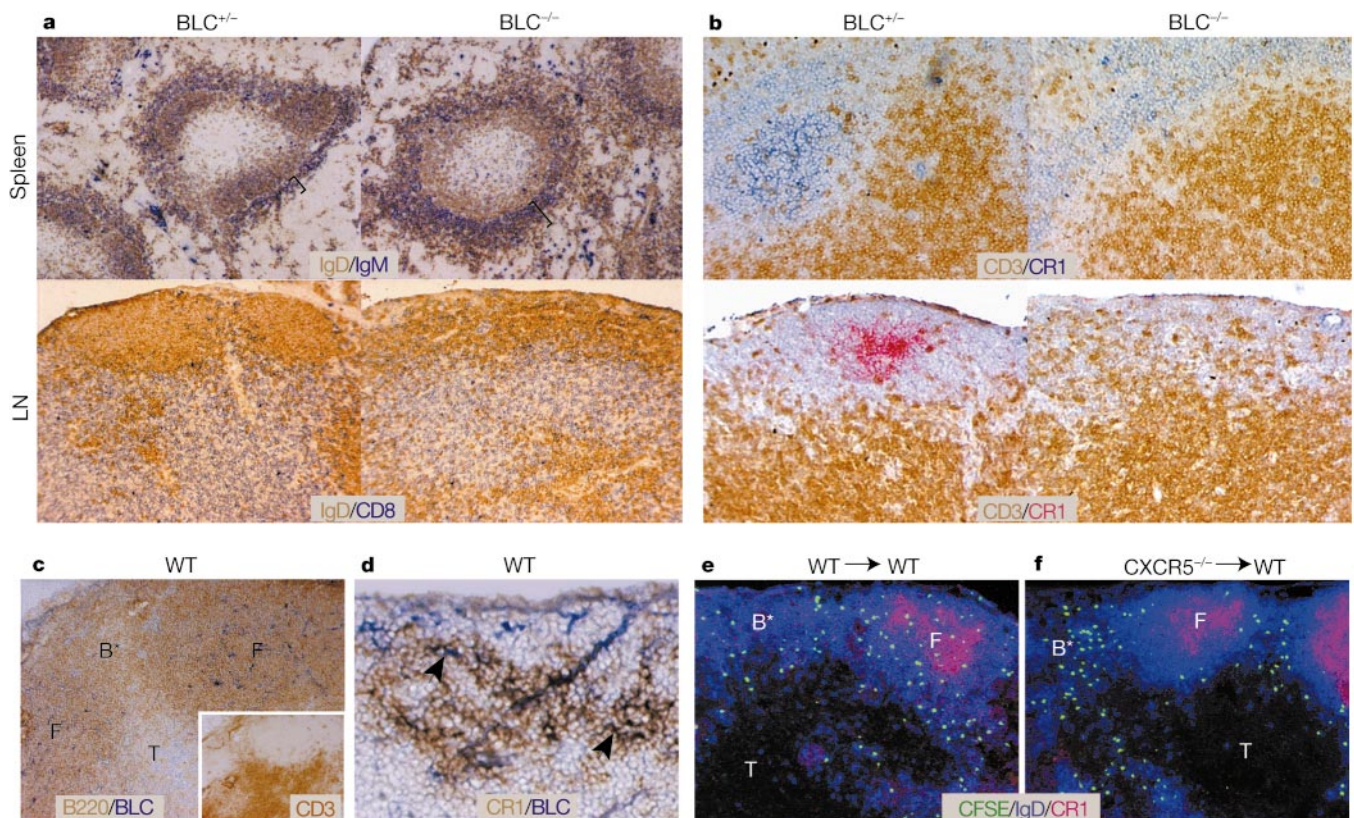


Figure 2 Absence of primary follicles and FDCs in BLC^{-/-} mice, and defective homing of CXCR5^{-/-} B cells to lymph node follicles. **a**, BLC^{+/-} and ^{-/-} spleen and lymph nodes stained for the indicated markers. Labels are in the same colour as the reaction product for that marker. Follicular B cells are mostly IgD^{hi}IgM^{lo} and appear brown, whereas marginal zone B cells are mostly IgD^{lo}IgM^{hi} and appear blue (top panels). Boundaries of the marginal zone are indicated by brackets. **b**, BLC^{+/-} and ^{-/-} spleen and lymph nodes stained to detect CR1 and T cells. FDCs appear as a network of intense staining for CR1, whereas splenic CR1^{hi} marginal zone B cells appear as a rim of faintly stained cells. Spleen and lymph nodes of more than 20 BLC^{+/-} and ^{-/-} mice were analysed for follicular

organization. **c**, Wild-type (WT) lymph node stained to detect BLC and B cells. Inset, adjacent section stained to detect T cells. **d**, High magnification view of a wild-type lymph node follicle stained to detect BLC and FDC (CR1). Arrows indicate examples of BLC and CR1 double-positive cells. Data are representative of four independent experiments. **e**, **f**, Distribution of transferred wild-type (**e**) or CXCR5^{-/-} (**f**) CFSE-labelled B cells in lymph nodes of wild-type recipients. Sections were stained to detect endogenous B cells and FDCs. Data are representative of four recipients. F, follicle; T, T zone; B*, BLC- and FDC-deficient area rich in B cells.

B-cell areas in response to BLC.

To test whether BLC directly induces LT α 1 β 2 expression, we incubated naive B cells *in vitro* with recombinant BLC (Fig. 4). BLC induced a dose-sensitive upregulation of LT α 1 β 2 on cultured B cells (Fig. 4a). The induction was inhibited by pre-incubation of the cells with PTX (Fig. 4b). Three other chemokines with chemotactic

activity on naive B cells were also tested: the broadly expressed CXCR4 ligand, SDF1; and SLC and ELC, two related CC chemokines made in the T-cell area that are ligands for CCR7 (ref. 19). Each of these chemokines caused detectable increases in LT α 1 β 2 expression on B cells, although in all cases the maximal induction was lower than the induction by BLC (Fig. 4c; and data not shown).

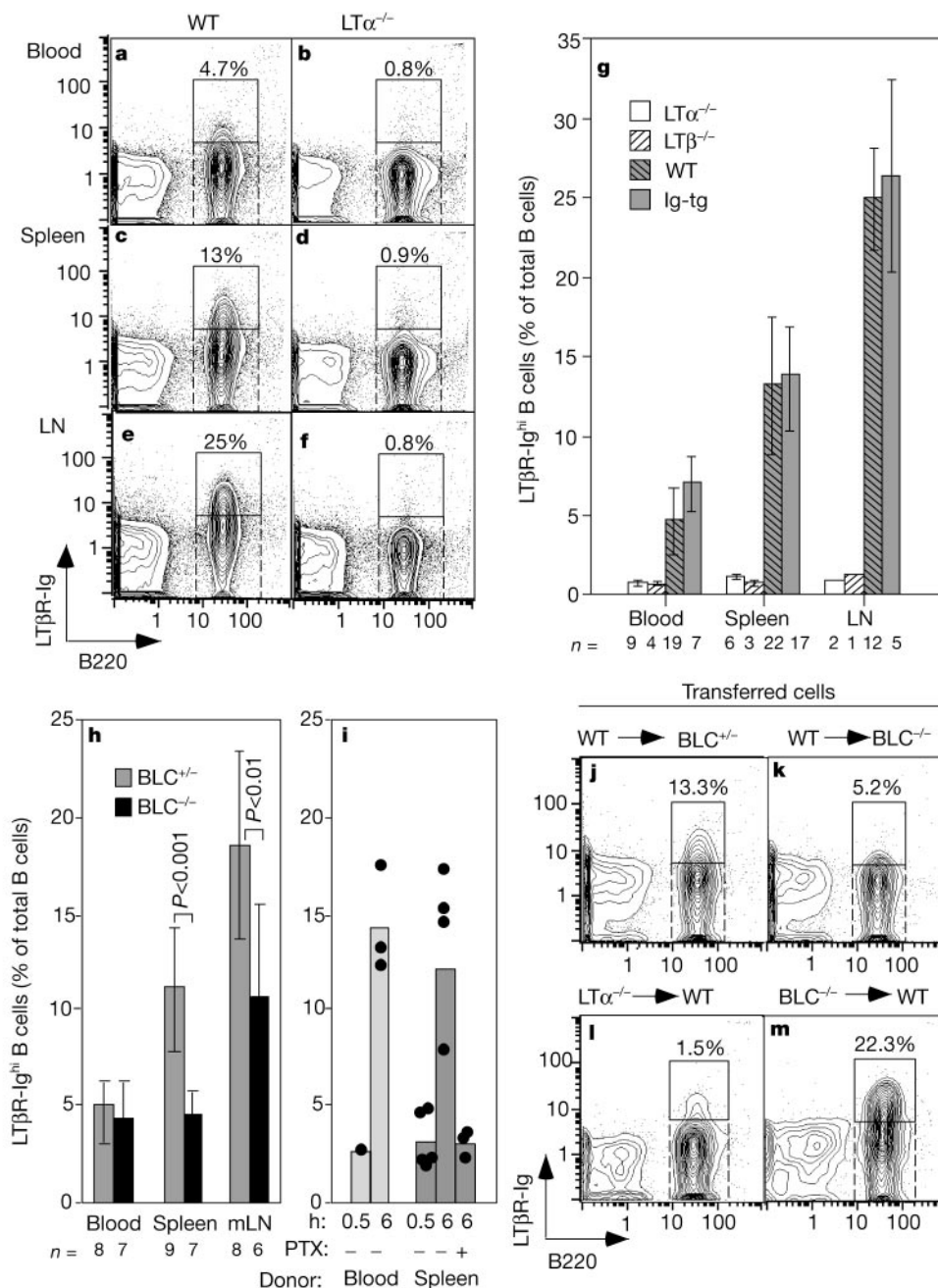


Figure 3 Membrane LT α 1 β 2 expression on B cells in peripheral lymphoid tissues and dependence on BLC. **a–f**, Flow cytometric analysis of cells from wild-type (left) and LT α ^{-/-} (right) blood (**a**, **b**), spleen (**c**, **d**) and lymph nodes (LN; **e**, **f**) stained for LT β R-Ig and B220. In **f**, lymph node cells were from irradiated wild-type mice that had been reconstituted for 8 weeks with LT α ^{-/-} bone marrow. Gates with solid lines show LT β R-Ig^{hi} B cells; total B cells were calculated as the cells in the region bounded by the dashed lines plus the cells within the LT β R-Ig^{hi} gate. Percentages of LT β R-Ig^{hi} B cells as a fraction of total B cells are shown above the windows. **g**, Summarized data from several flow cytometric analyses as shown in **a–f** with additional data obtained from LT β ^{-/-} and HEL-specific Ig-transgenic mice. **h**, LT α 1 β 2 expression levels on B cells from blood, spleen and mesenteric lymph nodes (mLN) of BLC^{-/-} animals compared with B cells of BLC^{+/-}

littermate controls. Cells were gated as in **a–f**. Bars represent the mean $\pm$ s.d. percentage of LT β R-Ig^{hi} B cells for the indicated number of animals. Statistical comparison is by Student's *t*-test. **i**, Upregulation of LT α 1 β 2 on B cells migrating into spleen occurs in a PTX-sensitive manner. Each point represents an individual recipient and bars represent the mean percentages of LT β R-Ig^{hi} B cells. **j–m**, LT α 1 β 2 upregulation on transferred B cells is dependent on BLC. Splenocyte donors and transfer recipients were as follows: Ly5.1⁺ wild type into BLC^{+/-} (**j**); Ly5.1⁺ wild type into BLC^{-/-} (**k**); Ly9.1⁺ LT α ^{-/-} into Ly9.2⁺ wild type (**l**); Ly9.1⁺ BLC^{-/-} into Ly9.2⁺ wild type (**m**). Data shown are LT β R-Ig and B220 staining of transferred Ly5.1⁺ (**j**, **k**) or Ly9.1⁺ (**l**, **m**) cells in recipient spleens 6 h after transfer, and are representative of 3 transfers of each type.

Although it is not yet clear whether these chemokines can induce LT α 1 β 2 expression on naive B cells *in vivo*, it seems possible that they contribute to the remaining expression in BLC $^{-/-}$ mice (Fig. 3h). BLC-mediated induction of LT α 1 β 2 was sensitive to actinomycin D, cycloheximide and wortmannin (Fig. 4d), providing evidence that BLC may signal through phosphatidylinositol-3-OH kinase to induce lymphotoxin transcription.

Despite low expression of LT α 1 β 2 on naive B cells and the

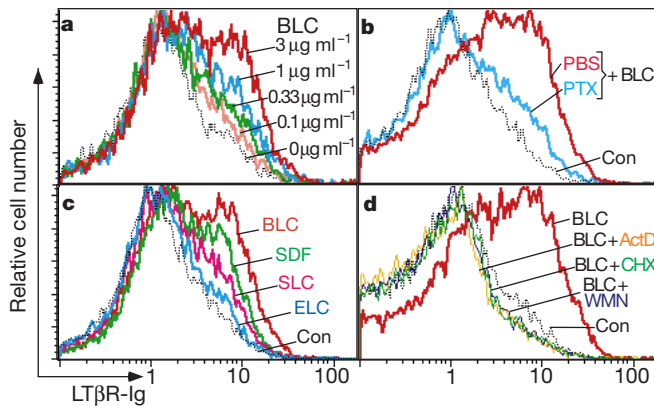


Figure 4 BLC induces LT α 1 β 2 upregulation on B cells *in vitro*. **a–d**, Histograms show LT β R-Ig staining of B220-gated B cells from *in vitro* cultures in the absence (Con) or presence of chemokines at the indicated amounts for 6 h. Cells were pre-incubated with 200 ng ml $^{-1}$ PTX or an equal volume of buffer for 2 h and then incubated with BLC (3 μ g ml $^{-1}$) for a further 6 h (**b**), incubated in SLC, ELC or BLC at 3 μ g ml $^{-1}$ or SDF1 at 1 μ g ml $^{-1}$ (**c**), or incubated with BLC (3 μ g ml $^{-1}$) alone or combined with actinomycin D (ActD, 5 μ g ml $^{-1}$), cycloheximide (CHX, 1 μ g ml $^{-1}$) or wortmannin (WMN, 100 nM) for 6 h (**d**).

absence of primary follicle FDCs in BLC-deficient mice, germinal centres formed in lymph nodes and spleen after immunization with a T-cell-dependent antigen (Fig. 5a). These germinal centres were misplaced and usually smaller than those found in wild-type controls, yet they contained CR1 $^{+}$ FDC networks (Fig. 5a). Similar observations for spleen germinal centre formation have been made in CXCR5-deficient mice²⁰. Germinal centre FDCs, like those of primary follicles, require LT α 1 β 2 for their development and maintenance^{4,21}. Staining with LT β R-Ig revealed high LT α 1 β 2 expression on GL7 $^{+}$ germinal centre B cells in wild-type mice (Fig. 5b). In contrast to our observations for naive B cells, LT α 1 β 2 expression on germinal centre B cells was not dependent on BLC (Fig. 5b). We conclude that the requirement for BLC to induce LT α 1 β 2 is overridden in germinal centres, most likely by T-cell-derived signals. CD40 signalling is necessary for germinal centre formation, and might provide the required signal for inducing LT α 1 β 2. In agreement with this possibility, CD40-mediated induction of LT α 1 β 2 is independent of CXCR5 (Fig. 5c).

In summary, our findings indicate that BLC has dual roles in follicular compartmentalization of B cells: mediating B-cell attraction, and inducing increased LT α 1 β 2 expression on the recruited cells. LT α 1 β 2 then engages LT β R on non-haematopoietic stromal cells, promoting maturation of FDCs and leading to increased expression of BLC^{4,5}. FDC maturation is regulated by signals in addition to LT α 1 β 2, including signals from tumour necrosis factor/tumour necrosis factor receptor-1 (ref. 4), and it remains to be determined whether these molecules function in the same or parallel pathways. In addition to showing the expression of LT α 1 β 2 on naive B cells, we have established the existence of a positive feedback loop between BLC and LT α 1 β 2. This feedback loop is likely to be critical in follicle development, causing the low amounts of BLC that are expressed independently of LT α 1 β 2 (ref. 5) to become upregulated as B cells are recruited. When the number of

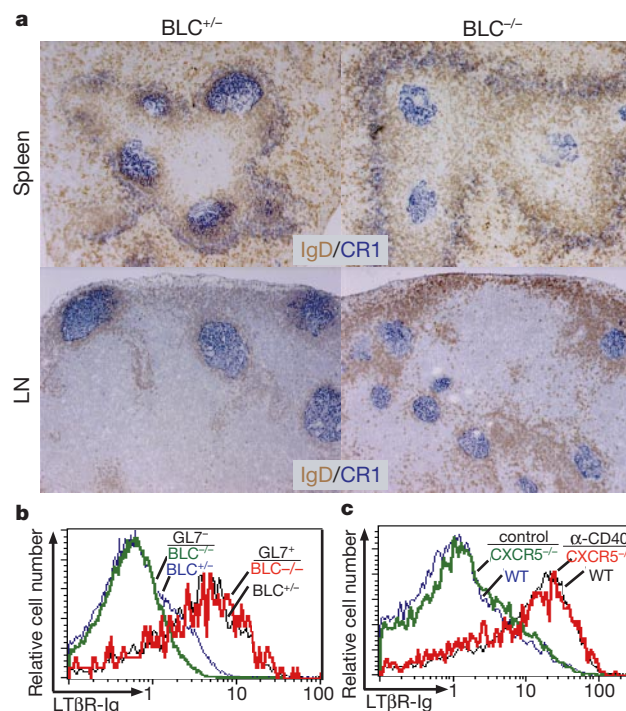


Figure 5 Misplaced FDC-containing germinal centres in BLC $^{-/-}$ mice and BLC-independent expression of LT α 1 β 2 on germinal centre B cells. **a**, Spleen and lymph nodes from immunized mice stained to detect non-germinal-centre B cells (IgD) and FDCs (CR1). Staining of serial sections with peanut agglutinin (PNA), established that the strongly CR1 $^{+}$ areas in the BLC $^{-/-}$ spleen and lymph node were germinal centres (not shown). Data are representative of eight immunized BLC $^{+/+}$ and BLC $^{-/-}$ mice. **b**, LT α 1 β 2

expression on germinal centre B cells. Histograms show LT β R-Ig staining on gated B220 $^{+}$ GL7 $^{-}$ non-germinal-centre and B220 $^{+}$ GL7 $^{+}$ germinal centre B cells from mesenteric lymph nodes of immunized BLC $^{+/+}$ and BLC $^{-/-}$ mice. **d**, *In vitro* CD40 stimulation induces LT α 1 β 2 expression in a CXCR5-independent manner. Spleen cells were incubated with buffer alone (control) or anti-CD40 antibody for 6 h and then stained. Histograms show LT β R-Ig staining on B220 $^{+}$ cells.

B cells entering an adult lymphoid organ increases, for example during an immune response, the feedback loop may help ensure that the follicular compartment can expand and accommodate the increased numbers of B cells. BLC-mediated B-cell recruitment and LT α 1 β 2 induction may also be involved in Peyer's patch organogenesis, because, in addition to requiring BLC and LT α 1 β 2, these structures depend strongly on B cells²². In support of this, ectopic expression of BLC induces B-cell-dependent and LT α 1 β 2-dependent lymphoid neogenesis²³.

A very early step in development of Peyer's patches and lymph nodes is the local accumulation of CD3⁺IL7R⁺ cells that express LT α 1 β 2 (refs 24–26). As some of these cells express CXCR5 (refs 25, 26), we propose that BLC functions in lymph node development by recruiting CXCR5⁺CD3⁺IL7R⁺ cells and inducing them to upregulate LT α 1 β 2. Finally, the finding that germinal centre B cells express LT α 1 β 2 and that this expression is independent of BLC, together with the detection of CR1⁺ FDCs in germinal centres of BLC^{−/−} mice, reveals that the pathway controlling FDC development in germinal centres is distinct from the BLC-mediated feedback loop operating in primary follicles. □

Methods

Generation of BLC^{−/−} mice

A 10-kilobase (kb) EcoRI fragment containing the *bcl* gene was isolated from a 129 mouse genomic DNA library (Genome Systems). A targeting vector was constructed in which nucleotides 18–116 of the second exon of *bcl* were replaced with an in-frame stop codon followed by a Mungo virus internal ribosomal entry site, the gene encoding enhanced green fluorescent protein and, in reverse orientation, a loxP-flanked neomycin resistance gene (*neo*^r). The linearized construct was electroporated into JM1 29 mouse ES cells (a gift from N. Killeen). G418 (200 μ g ml^{−1}) resistant colonies were screened by Southern hybridization of EcoRI-digested genomic DNA to a flanking 1.0-kb probe. Targeted clones were injected into C57BL/6 (B6) blastocysts. Chimaeric males were mated to B6 females, and germline transmission of the targeted allele was detected by Southern blot and by PCR using the primers: 5'-CGTCTATGTTCTTGTCCAATGGG-3' (sense); 5'-CTTACACAACCTCAGTTTGGGGC-3' (antisense; wild-type); and 5'-ACCTTGATTCCTTTGTGCGAGAGG-3' (antisense; mutant). Homozygous mutant mice were born at Mendelian ratios, were fertile and appeared healthy. Most mice used in this study were mixed 129 and B6 strain background carrying the original targeted allele, although the phenotype of these mice was indistinguishable from mice homozygous for a *neo*^r-deleted allele. Northern blot analysis was done as described⁵ using probes specific for *bcl* exon 2 or *EF-1 α* . BLC protein was detected in heparin-sepharose precipitates of 1% NP40 spleen homogenates by western blotting with polyclonal goat anti-mouse BLC (R&D Systems). B6 wild-type mice and some of the LT α ^{−/−} mice (on a mixed 129/B6 background²⁷) were from Jackson Labs. Other LT α ^{−/−} mice and the LT β ^{−/−} mice were from a B6 colony^{5,28}. Ig-transgenic mice were of the MD4 line¹⁵. To induce germinal centres, mice were injected intraperitoneally with 100 μ g of TNP- or NP-chicken γ -globulin (Biosearch Technologies) diluted in 200 μ l PBS and mixed with the Ribi Adjuvant System (Ribi Immunochem Research).

Flow cytometry, tissue staining and adoptive transfers

Surface expression of LT α 1 β 2 was detected with LT β R-Ig¹³, followed by anti-human IgG-PE (Jackson ImmunoResearch). Free IgG-binding sites were blocked by incubation with mouse and rat serum (1:25 dilution) for 15 min. Cells were then stained with appropriate antibody combinations for identifying lymphocyte subsets. Anti-LT β pre-blocking by incubation with 10 μ g ml^{−1} BBF6 (ref. 13) for 20 min inhibited LT β R-Ig staining of naive and germinal centre B cells. Flow cytometry, *in situ* hybridization, immunohistochemistry and immunofluorescence microscopy were done as described^{5,29}. For simultaneous detection of BLC and FDC, we incubated sections with goat anti-mBLC followed by donkey anti-goat IgG-biotin (Jackson). After colour development, sections were additionally stained with anti-CD35 (8C12). Anti-sheep IgG, anti-Armenian hamster IgG, anti-rat IgG-Cy3 and anti-rat IgG-HRP (Jackson), anti-rat IgG-biotin (Vector), anti-mouse IgD (Binding Site), anti-mouse IgM, anti-B220, anti-CD4 and anti-CD8 (Caltag), and all other antibodies (Pharmingen) were purchased, except anti-CD3 (2C11; a gift from L. Zuckerman). Biotinylated reagents were detected with streptavidin–Cychrome (flow cytometry; Pharmingen), Vectastain ABC-AP (immunohistochemistry; Vector) or streptavidin–AMCA (immunofluorescence; Vector). Adoptive transfers were done as described²⁹. In some experiments, donor cells were pre-incubated at 10⁷ ml^{−1} at 37 °C for 1 h with 200 ng ml^{−1} PTX (List Biological Laboratories). Spleenic B cells from CXCR5^{−/−} or wild-type donors were purified by staining non-B-cells with anti-CD43—biotin (Pharmingen), followed by MACS depletion and 5-(and-6)-carboxyfluorescein diacetate succinimidyl ester (CFSE) labelling as described²⁹.

Chemokines and cell culture

Recombinant mouse chemokines were produced as described²⁹. In preliminary experiments, *in vitro* incubation of B cells caused rapid upregulation of LT α 1 β 2, but after further

culture expression returned to levels similar to freshly isolated cells (data not shown). We therefore pre-incubated B6 spleen cells (10⁶ per well) in flat-bottom 96-well plates at 37 °C, 5% CO₂ in 200 μ l of RPMI/10% fetal calf serum for 14–15 h to allow adjustment to *in vitro* conditions. Chemokines or anti-CD40 (FGK; a gift from A. Rolink) was then added in 20 μ l of medium. In some experiments, PTX, actinomycin D (Calbiochem), cycloheximide (Calbiochem) or wortmannin (Sigma) was also added. After 6 h, cells were washed and analysed by flow cytometry. None of the chemokine preparations were found to cause upregulation of the activation marker CD69.

Received 18 February; accepted 31 May 2000.

1. Förster, R. *et al.* A putative chemokine receptor, BLR1, directs B-cell migration to defined lymphoid organs and specific anatomic compartments of the spleen. *Cell* **87**, 1037–1047 (1996).
2. Gunn, M. D. *et al.* A B-cell-homing chemokine made in lymphoid follicles activates Burkitt's lymphoma receptor-1. *Nature* **391**, 799–803 (1998).
3. Legler, D. F. *et al.* B cell-attracting chemokine 1, a human CXC chemokine expressed in lymphoid tissues, selectively attracts B lymphocytes via BLR1/CXCR5. *J. Exp. Med.* **187**, 655–660 (1998).
4. Fu, Y.-X. & Chaplin, D. D. Development and maturation of secondary lymphoid tissues. *Annu. Rev. Immunol.* **17**, 399–433 (1999).
5. Ngo, V. N. *et al.* Lymphotoxin- α/β and tumor necrosis factor are required for stromal cell expression of homing chemokines in B and T cell areas of the spleen. *J. Exp. Med.* **189**, 403–412 (1999).
6. Tew, J. G. *et al.* Follicular dendritic cells and presentation of antigen and costimulatory signals to B cells. *Immunol. Rev.* **156**, 39–52 (1997).
7. Fu, Y. X., Huang, G., Wang, Y. & Chaplin, D. D. B lymphocytes induce the formation of follicular dendritic cell clusters in a lymphotoxin α -dependent fashion. *J. Exp. Med.* **187**, 1009–1018 (1998).
8. Gonzalez, M., Mackay, F., Browning, J. L., Kosco-Vilbois, M. H. & Noelle, R. J. The sequential role of lymphotoxin and B cells in the development of splenic follicles. *J. Exp. Med.* **187**, 997–1007 (1998).
9. Endres, R. *et al.* Mature follicular dendritic cell networks depend on expression of lymphotoxin beta receptor by radioresistant stromal cells and of lymphotoxin beta and tumor necrosis factor by B cells. *J. Exp. Med.* **189**, 159–168 (1999).
10. Browning, J. L. *et al.* Lymphotoxin beta, a novel member of the TNF family that forms a heteromeric complex with lymphotoxin on the cell surface. *Cell* **72**, 847–856 (1993).
11. Worm, M. & Geha, R. S. CD40 ligation induces lymphotoxin alpha gene expression in human B cells. *Int. Immunol.* **6**, 1883–1890 (1994).
12. Pokholok, D. K. *et al.* Cloning and expression analysis of the murine lymphotoxin beta gene. *Proc. Natl Acad. Sci. USA* **92**, 674–678 (1995).
13. Browning, J. L. *et al.* Characterization of lymphotoxin-alpha beta complexes on the surface of mouse lymphocytes. *J. Immunol.* **159**, 3288–3298 (1997).
14. Miller, J. J. D., Johnsen, D. O. & Ada, G. L. Differences in localization of *Salmonella flagella* in lymph node follicles of germ-free and conventional rats. *Nature* **217**, 1059–1061 (1968).
15. Mason, D. Y., Jones, M. & Goodnow, C. C. Development and follicular localization of tolerant B lymphocytes in lysozyme/anti-lysozyme IgM/IgD transgenic mice. *Int. Immunol.* **4**, 163–175 (1992).
16. Gutman, G. A. & Weissman, I. L. Homing properties of thymus-independent follicular lymphocytes. *Transplantation* **16**, 621–629 (1973).
17. Kawabe, T. *et al.* The immune responses in CD40-deficient mice: impaired immunoglobulin class switching and germinal center formation. *Immunity* **1**, 167–178 (1994).
18. Vajdy, M., Kosco-Vilbois, M. H., Kopf, M., Kohler, G. & Lycke, N. Impaired mucosal immune responses in interleukin 4-targeted mice. *J. Exp. Med.* **181**, 41–53 (1995).
19. Cyster, J. G. Chemokines and cell migration in secondary lymphoid organs. *Science* **286**, 2098–2102 (1999).
20. Voigt, I. *et al.* CXCR5-deficient mice develop functional germinal centers in the splenic T cell zone. *Eur. J. Immunol.* **30**, 560–567 (2000).
21. Mackay, E. & Browning, J. L. Turning off follicular dendritic cells. *Nature* **395**, 26–27 (1998).
22. Golovkina, T. V., Shlomchik, M., Hannum, L. & Chervonsky, A. Organogenic role of B lymphocytes in mucosal immunity. *Science* **286**, 1965–1968 (1999).
23. Luther, S. A., Lopez, T., Bai, W., Hanahan, D. & Cyster, J. G. BLC expression in pancreatic islets causes B cell recruitment and lymphotoxin-dependent lymphoid neogenesis. *Immunity* **12**, 471–481 (2000).
24. Yokota, Y. *et al.* Development of peripheral lymphoid organs and natural killer cells depends on the helix-loop-helix inhibitor Id2. *Nature* **397**, 702–706 (1999).
25. Mebius, R. E., Rennett, P. & Weissman, I. L. Developing lymph nodes collect CD4⁺CD3⁺LT β ⁺ cells that can differentiate to APC, NK cells, and follicular cells but not T or B cells. *Immunity* **7**, 493–504 (1997).
26. Yoshida, H. *et al.* IL-7 receptor α ⁺ CD3⁺ cells in the embryonic intestine induces the organizing center of Peyer's patches. *Int. Immunol.* **11**, 643–655 (1999).
27. Matsumoto, M. *et al.* Role of lymphotoxin and the type I TNF receptor in the formation of germinal centres. *Science* **271**, 1289–1291 (1996).
28. Korner, H. *et al.* Distinct roles for lymphotoxin-alpha and tumor necrosis factor in organogenesis and spatial organization of lymphoid tissue. *Eur. J. Immunol.* **27**, 2600–2609 (1997).
29. Ansel, K. M., McHeyzer-Williams, L. J., Ngo, V. N., McHeyzer-Williams, M. G. & Cyster, J. G. *In vivo* activated CD4 T cells upregulate CXC chemokine receptor 5 and reprogram their response to lymphoid chemokines. *J. Exp. Med.* **190**, 1123–1134 (1999).
30. Tilney, N. L. Patterns of lymphatic drainage in the adult laboratory rat. *J. Anat.* **109**, 369–383 (1971).

Acknowledgements

We are especially grateful to N. Killeen for advice and help in generating the BLC knockout mice; K. Reif and L. Tang for helpful advice; and A. Schmidt and C. Backus for blastocyst transfers. K.M.A. is a HHMI predoctoral fellow, S.A.L. is supported by Human Frontier Science Program, and J.G.C. is a Packard fellow. This work was supported by grants from the NIH.

Correspondence and requests for materials should be addressed to J.G.C. (e-mail: cyster@itsa.ucsf.edu).

A tertiary interaction that links active-site domains to the 5' splice site of a group II intron

Marc Boudvillain^{†‡}, Alexandre de Lencastre^{*} & Anna Marie Pyle^{†‡}

^{*} Department of Biochemistry and Molecular Biophysics and [‡] Howard Hughes Medical Institute, Columbia University, New York, New York 10032, USA
[†] Present address: Centre de Biophysique Moléculaire, CNRS, Rue Charles Sadron, 45071 Orleans, Cedex 2, France.

Group II introns are self-splicing RNAs that are commonly found in the genes of plants, fungi, yeast and bacteria^{1,2}. Little is known about the tertiary structure of group II introns, which are among the largest natural ribozymes. The most conserved region of the intron is domain 5 (D5), which, together with domain 1 (D1), is required for all reactions catalysed by the intron³. Despite the importance of D5, its spatial relationship and tertiary contacts to other active-site constituents have remained obscure. Furthermore, D5 has never been placed directly at a site of catalysis by the intron. Here we show that a set of tertiary interactions (λ - λ') links catalytically essential regions of D5 and D1, creating the framework for an active-site and anchoring it at the 5' splice site. Highly conserved elements similar to components of the λ - λ' interaction are found in the eukaryotic spliceosome.

To dissect the structure of the group II intron active site, we developed a strategy based on two new chemo-genetic approaches: nucleotide analogue interference mapping (NAIM) and nucleotide analogue interference suppression (NAIS)⁴⁻⁹. Using the NAIM strategy, we identified the nucleotides and atoms most important for group II intron function⁵. The NAIM results were then used as a guide for designing group II intron mutations in regions most likely to affect active-site function. Intron mutants were screened and characterized kinetically to evaluate the mechanistic role and importance of each mutated nucleotide. We used a *trans*-branching assay (Fig. 1a)¹⁰ to explore the effects of mutations on transesterification and for optimizing conditions required for subsequent NAIS experiments⁵. For those mutants in which the branching reaction was strongly inhibited, we also conducted a 5'-splice-site hydrolysis assay^{11,12} to obtain the D5 dissociation constant (K_d) and the chemical rate of reaction (k_{chem}). In this way, we could evaluate the specific mutational effects on interdomain binding or on chemical catalysis. Among the variety of mutations that strongly affected function of exD123 RNA (M.B. and A.M.P., unpublished data), those with pronounced effects on reaction chemistry were concentrated in a region near the 5' end of D1 (represented by mutants exD123(G5:A) and exD123(A115:U), Table 1).

The partially functional exD123(G5:A) and exD123(A115:U)

mutants and the *trans*-branching reaction (Fig. 1) were used in NAIS experiments designed to reveal tertiary contacts between the mutated residues and atoms in D5 (ref. 5). Disruption of energetic coupling between the position of mutation and a given functional

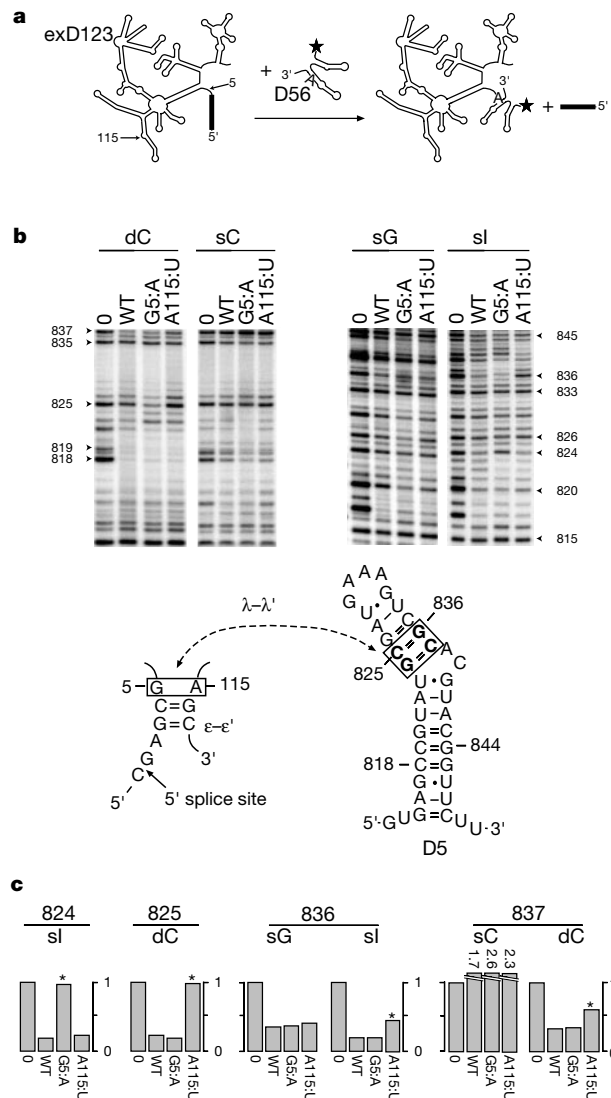


Figure 1 Nucleotide analogue interference suppressions within D5 and D1. **a**, The exD123 (5' exon/domains 1–3) mutants were tested for changes in the D5 interference pattern using the *trans*-branching assay. Branching results in the chemical ligation of D56 and D123, thereby ensuring a simple selection of reacted molecules. **b**, To search for NAIS signals within D5, we reacted wild-type (WT) or mutant exD123 transcripts with 5'-³²P-labelled D56 transcribed with an array of NTPαS analogues. Autoradiographs correspond to iodine-cleaved transcripts modified with guanosine-αS (sG), cytosine-αS (sC), deoxy-C-αS (dC) or inosine-αS (sl). Band intensities were quantitated and corrected for background phosphorothioate effects as described⁵. **c**, Normalized intensities of bands corresponding to positions where NAIS signals were observed (indicated by an asterisk). The normalized intensities are mean values, each having a maximum variance of $\pm 20\%$, from 2–4 independent experiments. Unlike the lower-intensity background bands, the NAIS signals are reproducible, dependent on iodine-cleavage and mutant-specific (**b**), as described⁵. Comparative data are shown for NαS-containing D56 that were unreacted (0) or branched to wild-type exD123 (WT), exD123(G5:A) (G5:A) and exD123(A115:U) (A115:U) mutants. For suppressions at positions 824 and 825, significant phosphorothioate effects were not observed and so corresponding bars are not shown. The band intensity observed for unreacted material was arbitrarily fixed to one. For branched products, the size of the bars is related to 1/(interference effect); that is, the smaller the bar, the stronger the interference. The values of band intensities over 1 are shown above the corresponding bars.

Table 1 Kinetic parameters of the exD123 mutants^{*}

exD123	Branching [†] k_{obs} (10^{-4} min^{-1})	Hydrolysis k_{chem} (10^{-3} min^{-1})	K_d (μM)
Wild type	250 $\pm$ 10	87 $\pm$ 4	0.30 $\pm$ 0.04
G5:A	1.8 $\pm$ 0.2	8 $\pm$ 3	0.27 $\pm$ 0.09
A115:U	2.4 $\pm$ 0.1	2.7 $\pm$ 0.2	0.23 $\pm$ 0.09
A198:U [‡]	29 $\pm$ 2	52 $\pm$ 7	3.6 $\pm$ 0.9

^{*} Branching and hydrolysis reactions were performed as described in Methods.

[†] For exD123(G5:A) and exD123(A115:U), the reaction end point could not be determined accurately and was assumed to be 0.44 as for wild-type exD123. For exD123(A198:U), the reaction end point is 0.12.

[‡] A mutation that induces binding defects is presented for comparison: the A198→U mutation disrupts a ground-state hydrogen bond in the known ζ - ζ' interaction between D1 and D5 (ref. 16), as inferred from the crystal structure of an analogous GNRA tetraloop-receptor complex²⁷. The 10-fold increase in K_d corresponds to the loss of $\sim 1.6 \text{ kcal mol}^{-1}$ of D1–D5 interaction energy, which is consistent with the loss of at least one hydrogen bond.

group was revealed by suppression of an interference normally observed in D56 upon nucleotide analogue substitution (Fig. 1). The pattern of D56 interferences obtained with wild-type exD123 is similar to that previously described⁵; however, several differences were observed with the exD123(G5:A) and exD123(A115:U) mutants (Fig. 1). The G5→A mutation in D1 induces complete suppression of inosine interference at G824 in D5, whereas the A115→U mutation suppresses deoxynucleotide interference at C825. In addition, the exD123(A115:U) mutant suppresses an inosine interference at G836 and partially rescues a strong deoxynucleotide interference at C837.

These results show energetic coupling between two nucleotides of D1 (G5 and A115) and two consecutive G-C pairs located at the base of D5 stem 2. This energetic coupling is likely to reflect specific interactions because the exD123(G5:A) and exD123(A115:U) mutants, which have similar reactivities (Table 1), yield NAIS signals that are not observed upon mutation of other positions in D1 or D3 (Fig. 1)⁵. The new interaction, which we have named λ - λ' , is composed of nucleotides that exhibit a high level of

phylogenetic conservation and which show strong effects when mutated (M.B., A.D.L. and A.M.P., unpublished data).

Notably, the adjacent ϵ - ϵ' interaction¹³ brings G5 and A115 into close proximity (Fig. 1). The ϵ - ϵ' motif contributes to 5'-splice-site selection¹³, and it has also been observed that the ϵ -nucleotides G3 and C4 must be on the same molecule as their partners G116 and C117 for optimal chemical reactivity of ribozymes derived from the intron (S. Leuvin and A.M.P., unpublished data). This suggests that ϵ - ϵ' itself may be a component of the intron active-site. The λ - λ' interaction therefore serves to join the ϵ - ϵ' nucleotides in D1 with catalytically essential functionalities in D5 (refs 12, 14, 15), a role that is consistent with the fact that λ - λ' mutations have marked effects on reaction chemistry (Table 1).

The λ - λ' interaction is the third contact identified between D5 and another domain of the intron. The first contact identified (ζ - ζ')¹⁶ promotes D1-D5 association through a tetraloop-receptor interaction motif. Mutation of ζ - ζ' functionalities induces binding defects in activity assays (Table 1)¹². The κ - κ' interaction also contributes primarily to the ground-state binding of D1 to D5 (M.B. and A.M.P., unpublished data). Consistent with this, both ζ - ζ' and κ - κ' involve atoms that lie on the so-called 'binding face' of D5 (refs 12, 15). Conversely, λ - λ' involves D1 and D5 function-

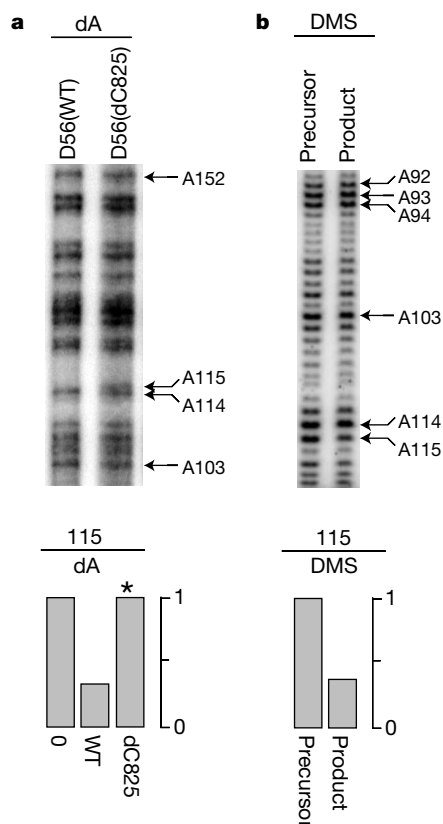


Figure 2 Single-atom NAIS and D5 interferences at A115. **a**, To search for NAIS signals within D1, we reacted wild-type exD123 transcripts containing NTP α S with 5'-³²P-labelled D56 molecules containing a single deoxynucleotide at C825 (D56(dC825)). The interference suppression observed for dA115 is specific, that is, it is not observed in the presence of other D56 modifications (such as dC837), and dC825 does not cause a suppression of dG5 interference (see Supplementary Information). Note that in these experiments, the 5' exon was only 17 nucleotides in length (17D123, ref. 5). **b**, The involvement of A115 N1 is implicated by a D5-dependent DMS interference at position A115. Folded exD123 was reacted with DMS (1:100 dilution) for 5 min at room temperature. The DMS-modified exD123 was incubated with 3 μ M D5 at 45 °C for 20 min under standard hydrolysis conditions¹¹. Precursor exD123 was separated from the D123 product by PAGE, eluted and then primer extended using AMV reverse transcriptase for 15 min at 55 °C. After PAGE, products were analysed and quantitated on a PhosphorImager (Molecular Dynamics). The A115 interference value shown represents the average from three experiments, with a variance of 18%.

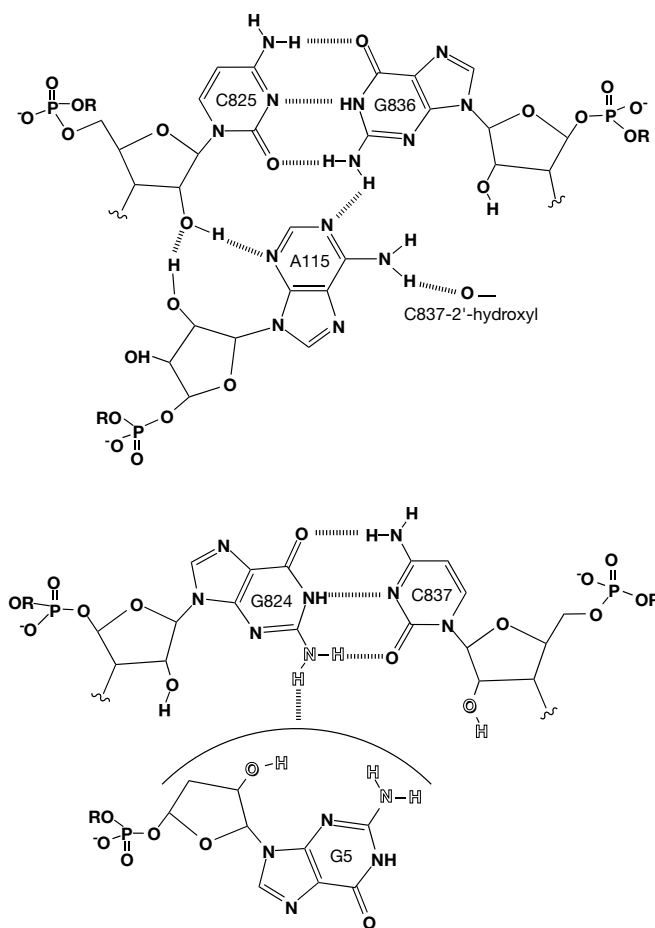


Figure 3 Model for the λ - λ' interaction. Proposed arrangements of the two minor groove triples formed by A115 and G5 with the (5'-GC)-(5'-GC) tandem pairs in stem 2 of D5. Modelling (Insight II, Biosym/MSI) of the stacked triple interactions did not reveal any significant steric or geometric problems with the arrangements shown. Owing to a lack of constraints, the G824 amine is shown interacting with an undefined group on G5. Functionalities in this triple that have been implicated by NAIM experiments are shown in open lettering, indicating potential sites of interaction within the triple or with other nucleotides.

alities that have specific roles in catalysis (Table 1)¹². It is therefore striking that the λ - λ' interaction occurs on the previously defined 'chemical face' of D5 (refs 12, 15), where catalysis probably occurs.

All bonds in the proposed model for λ - λ' involve atoms in the minor groove of both D5 and D1 components. The suppression data show that G5 interacts with N2 of G824, and that A115 interacts with N2 of G836 and the 2'-OH of C837 and C825. Previous work showed that G5 interacts through 2'-OH and N2 functionalities⁵, whereas A115 interacts through a 2'-OH group⁵. Evidence for direct contact between the N1 of A115 and elements of D5 is provided by D5-dependent dimethyl sulphate (DMS) footprint¹⁷ and interference signals at A115 (Fig. 2). Further evidence for a hydrogen bond between C825 and D1 was obtained through an adaptation of the single-atom NAIS approach⁴⁻⁹, in which a chimaeric D56 molecule containing a single deoxynucleotide at C825 was reacted with exD123 transcripts modified with NTP α S. In that case, the specific suppression of a deoxynucleotide interference was observed at A115 (Fig. 2, and see Supplementary Information). The spatial arrangement that agrees best with this data is one in which the minor groove edge of A115 forms a triple interaction in the minor groove of the C825-G836 base pair (Fig. 3) using an arrangement that has been observed crystallographically¹⁸. G5 is likely to form a second minor groove triple with the G824-C837 base pair (Fig. 3). Because of the limited number of available constraints, this triple interaction is represented in a more general form.

On the basis of this information, the λ - λ' interaction can be visualized as an extended minor-groove triplex that orientates the ϵ - ϵ' helix along the 'chemical face' of D5, placing ϵ - ϵ' against the D5 bulge (Fig. 4). A composite ϵ - ϵ' / λ - λ' interaction that involves backbone-backbone interactions with the D5 bulge is inferred by strong deoxy and phosphorothioate interferences in ϵ - ϵ' and the D5 bulge^{5,12,19}, as well as by evidence for metal ion binding in this region (J. Swisher, R. Sigel and A.M.P., unpublished data). A composite ϵ - ϵ' / λ - λ' interaction (Fig. 4) would align the 5' splice site next to major groove functionalities of G817 that have been shown to be critical for chemical catalysis by group II introns¹⁵. Thus, λ - λ' is a novel substructure that links active-site constituents together and anchors them at a site of group II intron reaction.

The λ - λ' interaction involves the fifth intronic nucleotide, which is invariably a guanine in group II introns²⁰ and in most nuclear introns²¹. In spliceosomal introns, G5 is part of a three-base-pair

helix between the first several nucleotides of the intron and a section of U6 RNA called the 'ACAGA-box' (paired residues are underlined)^{22,23}, forming a duplex that may be analogous to the ϵ - ϵ' interaction. It will be interesting to test whether this region interacts with a set of universally conserved G-C pairs in helix 2 of the U2-U6 complex, which has been proposed to be analogous to stem 2 of D5 (refs 19, 24, 25). The existence of such a spliceosomal λ - λ' homologue might strengthen the hypothesis of an evolutionary relationship between active sites involved in nuclear pre-mRNA splicing and group II introns. □

Methods

Transcriptions

Plasmids encoding the mutant transcripts were prepared from precursor plasmid pJDI3'-673 (ref. 26) using the QuikChange site-directed mutagenesis kit (Stratagene) or according to described procedures⁵. Plasmids derived from pJDI3'-673 (encoding exD123 RNA) were linearized with *Bam*HI, pT7-D56 (encoding D56) with *Eco*RV (ref. 10) and pJDI5'-75 (encoding D5) with *Hpa*II (ref. 11) and transcribed as described^{10,11}. Transcripts randomly modified with NTP α S analogues were prepared as described⁵.

Reaction kinetics

Hydrolysis and *trans*-branching kinetic assays were carried out as described^{10,11} with minor modifications. For branching kinetics, 5'-³²P end-labelled D56 transcripts (10 nM, final concentration) were mixed with exD123 (3 μ M, final concentration) and heated at 95 °C for 1 min. We cooled the mixture to 42 °C and added salts (100 mM MgCl₂, 0.5 M NH₄Cl, 40 mM MOPS, pH 6) before further incubation at 42 °C. At various times, aliquots were withdrawn, quenched as described⁵ and subjected to electrophoresis on stacked denaturing 4%:20% (top:bottom) polyacrylamide gels (PAGE). Hydrolysis kinetics were carried out similarly except that 5'-³²P end-labelled exD123 transcripts (1-5 nM, final concentration) were incubated with increasing concentrations of D5 (50 nM to 10 μ M, final concentration) in 100 mM MgCl₂, 0.5 M KCl, 40 mM MOPS, pH 7.5 at 45 °C. Reactions were quantitated as described¹¹.

Nucleotide analogue interference suppression (NAIS)

We mixed wild-type exD123 (0.5 μ M final concentration) and exD123 mutants (3 μ M final concentration) with 5'-³²P end-labelled D56 transcripts (10 nM final concentration) that were randomly modified with a specific NTP α S analogue. The mixture was pre-incubated as described above and then reacted at 42 °C in 100 mM MgCl₂, 2 mM Mn(AcO)₂, 0.5 M (NH₄)₂SO₄, 40 mM MOPS, pH 6 (ref. 5). After 3 h of incubation at 42 °C, reaction products were resolved by PAGE. Iodine cleavage was performed as described⁵, and the products were resolved by PAGE. Gel bands were quantitated using a phosphorimager (Molecular Dynamics) as described⁵. For single-atom NAIS, D56 molecules containing a deoxynucleotide at 825 (obtained by ligation of two synthetic oligonucleotides) were 5'-³²P end-labelled and were reacted with exD123 mutants (3 μ M final concentration) randomly modified with NTP α S analogues under the conditions described above. Analysis of interferences in D1 was done as described⁵.

Received 13 January; accepted 5 April 2000.

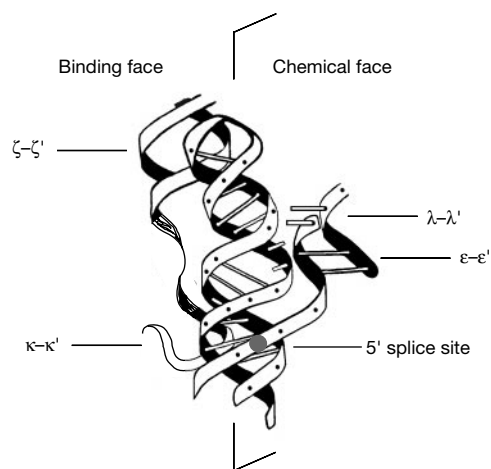


Figure 4 Representation of the putative group II intron active site, showing regions of D1 interacting with D5. On the left side of the model are the ζ - ζ' and κ - κ' contacts^{5,16} between the D5 binding face¹² and the tetraloop receptor and three-way junction of D1. The right side of the model shows the D5 chemical face interacting with the extended λ - λ' / ϵ - ϵ' interactions that align the 5' splice site over the D5 catalytic locus¹⁵.

- Michel, F. & Ferat, J.-L. Structure and activities of group II introns. *Annu. Rev. Biochem.* **64**, 435-461 (1995).
- Pyle, A. M. in *Nucleic Acids and Molecular Biology* (eds Eckstein, F. & Lilley, D. M. J.) 75-107 (Springer, New York, 1996).
- Qin, P. & Pyle, A. The architectural organization and mechanistic function of group II intron structural elements. *Curr. Opin. Struct. Biol.* **8**, 301-308 (1998).
- Strobel, S. A. & Shetty, K. Defining the chemical groups essential for *Tetrahymena* group I intron function by nucleotide analog interference mapping. *Proc. Natl Acad. Sci. USA* **94**, 2903-2908 (1997).
- Boudvillain, M. & Pyle, A. Defining functional groups, core structural features and inter-domain tertiary contacts essential for group II intron self-splicing: a NAIM analysis. *EMBO J.* **17**, 7091-7104 (1998).
- Ortoleva-Donnelly, L., Swczak, A. A., Gutell, R. R., Strobel, S. A. The chemical basis of adenosine conservation through the *Tetrahymena* ribozyme. *RNA* **4**, 498-519 (1998).
- Ortoleva-Donnelly, L., Kronman, M., Strobel, S. A. Identifying RNA minor groove tertiary contacts by nucleotide analogue interference mapping with N2-methylguanosine. *Biochemistry* **37**, 12933-12942 (1998).
- Strobel, S. A., Ortoleva-Donnelly, L., Ryder, S. P., Cate, J. H., Moncoeur, E. Complementary sets of noncanonical base pairs mediate RNA helix packing in the group I intron active site. *Nature Struct. Biol.* **5**, 60-65 (1998).
- Swczak, A., Ortoleva-Donnelly, L., Ryder, S. P., Moncoeur, E. & Strobel, S. A. A minor groove RNA triple helix within the catalytic core of a group I intron. *Nature Struct. Biol.* **5**, 1037-1042 (1998).
- Chin, K. & Pyle, A. M. Branchpoint attack in group II introns is a highly reversible transesterification, providing a potential proofreading mechanism for 5' splice site selection. *RNA* **1**, 391-406 (1995).
- Pyle, A. M. & Green, J. B. Building a kinetic framework for group II intron ribozyme activity: quantitation of interdomain binding and reaction rate. *Biochemistry* **33**, 2716-2725 (1994).
- Abramovitz, D. L., Friedman, R. A. & Pyle, A. M. Catalytic role of 2'-hydroxyl groups within a group II intron active site. *Science* **271**, 1410-1413 (1996).
- Jacquier, A. & Michel, F. Base-pairing interactions involving the 5' and 3' terminal nucleotides of

- group II self-splicing introns. *J. Mol. Biol.* **213**, 437–447 (1990).
14. Peebles, C. L., Zhang, M., Perlman, P. S. & Franzen, J. F. Identification of a catalytically critical trinucleotide in domain 5 of a group II intron. *Proc. Natl Acad. Sci. USA* **92**, 4422–4426 (1995).
 15. Konforti, B. B. *et al.* Ribozyme catalysis from the major groove of group II intron domain 5. *Mol. Cell* **1**, 1–20 (1998).
 16. Costa, M. & Michel, F. Frequent use of the same tertiary motif by self-folding RNAs. *EMBO J.* **14**, 1276–1285 (1995).
 17. Costa, M. & Michel, F. Tight binding of the 5' exon to domain I of the group II self-splicing intron requires completion of the intron active sites. *EMBO J.* **18**, 1025–1037 (1999).
 18. Pley, H. M., Flaherty, K. M. & McKay, D. B. Model for an RNA tertiary interaction from the structure of an intermolecular complex between a GAAA tetraloop and an RNA helix. *Nature* **372**, 111–113 (1994).
 19. Chanfreau, G. & Jacquier, A. Catalytic site components common to both splicing steps of a group II intron. *Science* **266**, 1383–1387 (1994).
 20. Michel, F., Umesono, K. & Ozeki, H. Comparative and functional anatomy of group II catalytic introns—a review. *Gene* **82**, 5–30 (1989).
 21. Moore, M. J., Query, C. C. & Sharp, P. A. In *The RNA World* (eds Gesteland, R. F. & Atkins, J. F.) 303–358 (Cold Spring Harbor Laboratory Press, New York, 1993).
 22. Kandels-Lewis, S. & Seraphin, B. Involvement of U6 snRNA in 5' splice site selection. *Science* **262**, 2035–2039 (1993).
 23. Lesser, C. F. & Guthrie, C. Mutations in U6 snRNA that alter splice site specificity: implications for the active site. *Science* **262**, 1982–1988 (1993).
 24. Madhani, H. D. & Guthrie, C. A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosomes. *Cell* **71**, 803–817 (1992).
 25. Yu, Y. T., Maroney, P. A., Darzynkiewicz, E. & Nilsen, T. W. U6 snRNA function in nuclear pre-mRNA splicing: a phosphorothioate interference analysis of the U6 phosphate backbone. *RNA* **1**, 46–54 (1995).
 26. Jarrell, K. A., Dietrich, R. C. & Perlman, P. S. Group II intron domain 5 facilitates a trans-splicing reaction. *Mol. Cell. Biol.* **8**, 2361–2366 (1988).
 27. Cate, J. H. *et al.* Crystal structure of a group I ribozyme domain: Principles of RNA packing. *Science* **273**, 1678–1685 (1996).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank Q. Liu for preparing plasmids encoding exD123 mutants. We also thank S. Strobel for helpful discussions, R. Sousa for the clone of the Y639F RNA polymerase and P. S. Perlman for the gift of plasmids pJDI5'-75 and pJDI3'-673. M.B. is a Postdoctoral Research Associate and A.M.P. is an Assistant Investigator with the Howard Hughes Medical Institute, which we thank for financial support of this work. A.D.L. is supported by a fellowship from PRAXIS XXI (Portugal).

Correspondence and requests for material should be addressed to A.M.P. (e-mail: amp11@columbia.edu).

A ratchet-like inter-subunit reorganization of the ribosome during translocation

Joachim Frank*†‡ & Rajendra Kumar Agrawal†‡

* Howard Hughes Medical Institute, † Health Research Incorporated at the Wadsworth Center, and ‡ Department of Biomedical Sciences, State University of New York at Albany, Empire State Plaza, PO Box 509, Albany, New York 12201-0509, USA

The ribosome is a macromolecular assembly that is responsible for protein biosynthesis following genetic instructions in all organisms. It is composed of two unequal subunits: the smaller subunit binds messenger RNA and the anticodon end of transfer RNAs, and helps to decode the mRNA; and the larger subunit interacts with the amino-acid-carrying end of tRNAs and catalyses the formation of the peptide bonds. After peptide-bond formation, elongation factor G (EF-G) binds to the ribosome, triggering the translocation of peptidyl-tRNA from its aminoacyl site to the peptidyl site, and movement of mRNA by one codon¹. Here we analyse three-dimensional cryo-electron microscopy maps of the *Escherichia coli* 70S ribosome in various functional states, and show that both EF-G binding and subsequent GTP

hydrolysis lead to ratchet-like rotations of the small 30S subunit relative to the large 50S subunit. Furthermore, our finding indicates a two-step mechanism of translocation: first, relative rotation of the subunits and opening of the mRNA channel following binding of GTP to EF-G; and second, advance of the mRNA/(tRNA)₂ complex in the direction of the rotation of the 30S subunit, following GTP hydrolysis.

We analysed the amplitude-corrected three-dimensional cryo-electron microscopy (3D cryo-EM) density maps of five complexes presenting the 70S ribosome in different functional states: (1) an initiation-state complex², bound with a mRNA and fMet-tRNA^{fMet} at the peptidyl (P) site (the 'control'); (2) the pre-translocational state complex³, bound with mRNA, aminoacyl (A)- and P-site tRNAs, and EF-G in the presence of a nonhydrolysable GTP analogue, GMPP(CH₂)P (the 'GTP form'); (3) as (2), but without mRNA and tRNAs³; (4) the post-translocational state complex³, bound with mRNA, P and exit (E)-site tRNAs, and EF-G in the presence of GTP and fusidic acid (the 'GDP form'); and (5) as (4), but without mRNA and tRNAs^{3,4}. The comparison of these refined maps shows that, upon binding of EF-G/GMPP(CH₂)P to the ribosome, the 30S subunit rotates with respect to the 50S subunit by an angle of about 6 (± 0.7) degrees, resulting in a maximum displacement, at the periphery of the ribosome, of 19 (± 3) Å. The rotation is counterclockwise when seen from the solvent side of the 30S subunit (compare Fig. 1a and b). Following EF-G-dependent GTP hydrolysis, in the presence of fusidic acid the 30S subunit rotates back (Fig. 1c) by an angle of about 3 (± 0.7) degrees. The conformational rearrangement is marked if one compares the distances between the L1-protein region of the 50S subunit and the platform and head regions of the 30S subunit (Fig. 1a and b). The positions of beak and spur of the 30S subunit are also strongly affected. Figure 1 shows maps of 70S/EF-G complexes lacking mRNA and tRNAs. The lower occupancy of EF-G/GMPP(CH₂)P in the complex containing tRNA at the A site^{3,5} makes its presentation confusing, however the rotations of the subunit in the maps with or without mRNA and tRNAs are of the same size.

The rotation is accompanied by many conformational changes of the two subunits. Among these is a bifurcation of the stalk of the 50S subunit in the GTP state^{3,5} (Fig. 1b) and its reversal to a single, elongated form in the GDP state^{3–5} (Fig. 1c), showing an arc-like connection between the G' domain of EF-G and the stalk base.

Going from the control to the GTP state, we also see a transition to a less compact structure⁶, such that the radius of gyration increases⁷ by 1.1 Å, and a complex rearrangement in the multi-stranded mass near the mRNA channel on the 30S subunit^{8–11}. Significantly, both the entrance channel (conducting the downstream 3' end of the mRNA) and the exit channel (conducting the upstream 5' end of the mRNA) are affected (Fig. 2). The entrance channel is defined by a contact between the 30S subunit shoulder and head^{12,13} involving the 530 loop, which is dynamic during decoding¹⁴, and the 1050/1200 regions of the 16S rRNA¹⁵, respectively. The entrance channel is much wider in both GTP (an increase in cross-sectional area by ~68%) and GDP (by ~118%) states than in the control (Fig. 3). Thus, the opening at the entrance site is wider (by ~30%) in the GDP (Figs 2b and 3c) than in the GTP (Figs 2a and 3b) state. The shape and position of the exit channel, formed between the platform and the head¹¹, are known to be variable^{12,13}. The behaviour of this channel is the reverse of that of the entrance channel: it is wider in the GTP (Fig. 2c) than in the GDP state (Fig. 2d). Thus, a dynamic clamping mechanism apparently exists that alternatively secures the mRNA in the entrance channel or leaves it free to move following EF-G/GTP binding and GTP hydrolysis.

We find clear evidence in our cryo-maps that the positions of the tRNAs with respect to the 30S subunit are unchanged in the GTP state (Fig. 3b; ref. 16), indicating that translocation of the tRNA from the A to the P site has not taken place³. However, a contact

between the head and the anticodon arm of the A-site tRNA (formed before EF-G binding) is severed (arrow, Fig. 3b), which would be required as a preparation for the translocation of tRNA. In addition, the subunit head flattens from a more convex shape and gains in length on the platform side. After GTP hydrolysis (Fig. 3c), the A site is vacated and the P and E sites are occupied. *In vitro*, the E site is always occupied by deacylated tRNA¹⁷, and the corresponding mass is visible in Fig. 3b and c. However, the density corresponding to the E-site tRNA is higher after GTP hydrolysis (not shown). The conformation of the upper half of the helix 44 of the 16S RNA^{18–20}, which is clearly distinguishable in all maps (Fig. 3) and the tip of which forms the decoding region, is affected by both EF-G binding and GTP hydrolysis.

These observations allow us to sketch out the following sequence of events (Fig. 4): EF-G binding alone triggers a relative movement of the 30S versus the 50S subunit accompanied by a widening of the mRNA channel, without changing the position of the tRNA ends relative to the 30S or 50S subunit, although

inevitably leading to a skewed tRNA orientation. Subsequent GTP hydrolysis triggers conformational changes that result in three coupled events: (1) translocation of the mRNA/(tRNA)₂ complex with respect to the small subunit, which requires the mRNA to be free to move; (2) release of EF-G; and (3) reverse rotation toward the original positions of the subunits, a movement that is arrested in an intermediate position owing to the presence of fusidic acid^{3,4}.

Thus, the translocation process entails two relative rotations of the subunits in opposite directions, which are rapid on the timescale of the elongation cycle, and can only be observed by trapping intermediate states. The observation of a rotation preceding the actual translocation calls for a more detailed terminology in describing ribosome states (Fig. 4): 'pre', before EF-G binding; 'pre'', EF-G/GTP bound and subunits rotated, but before GTP hydrolysis; 'post', with EF-G/GDP bound but immediately after GTP hydrolysis; 'post^{fus}', EF-G/GDP bound in the presence of fusidic acid; and 'post', after release of EF-G and the subunits have

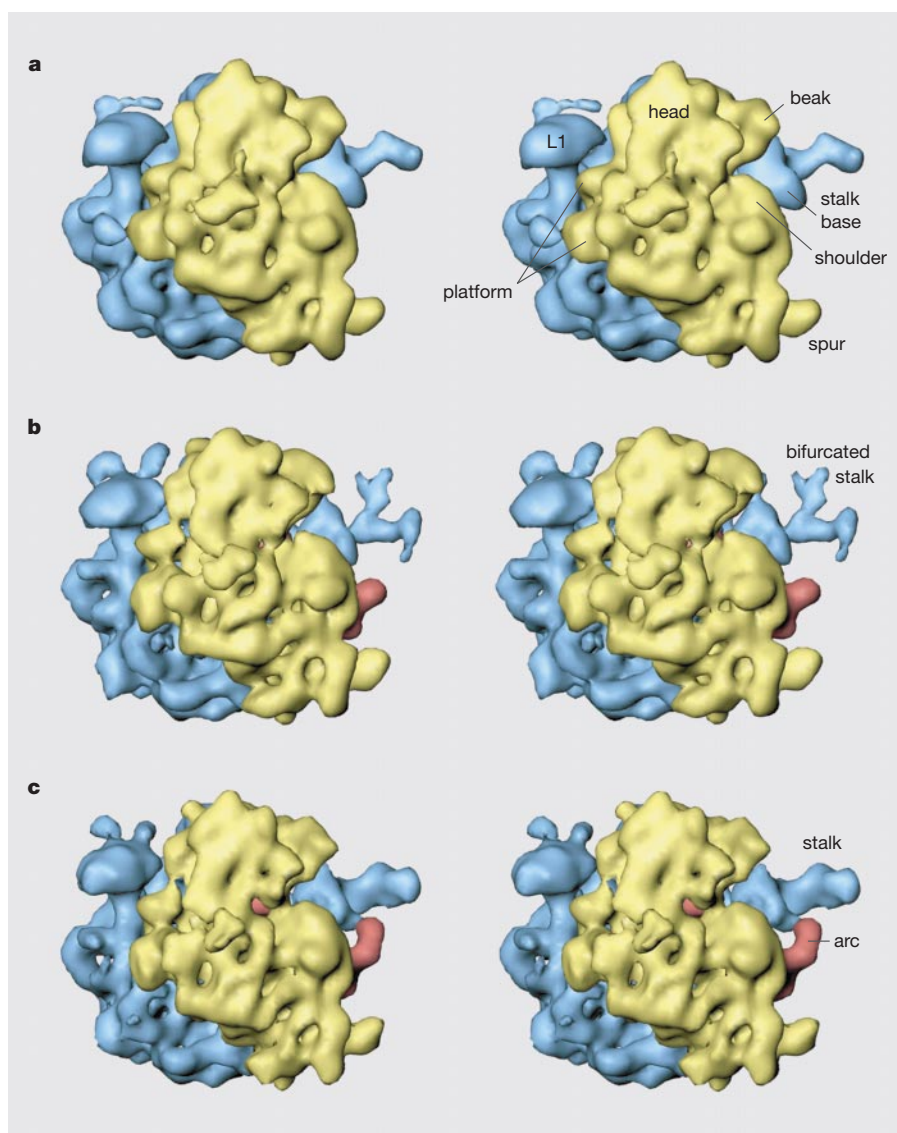


Figure 1 Stereo representation of the rotational movement of the 30S subunit (yellow) with respect to 50S subunit (blue) upon EF-G (red) binding. **a**, 70S/fMet-tRNA^{Met} complex; **b**, 70S/EF-G/GMPP(CH₂)P complex; and **c**, 70S/EF-G/GDP/fusidic acid complex. The three-dimensional maps of the 70S complexes are shown from the 30S

solvent side. The masses corresponding to EF-G in **b** and **c** are only partially seen in this view. The movement of the 30S subunit is anticlockwise from **a** to **b** and clockwise from **b** to **c**. Landmarks: L1, protein L1; arc, arc-like connection between EF-G and stalk base.

rotated into their original relative position. The orientation between the subunits is the same in the pre and post states^{17,21}. The same relative orientations of the subunits in both states would be required for the binding of EF-G and the highly similar EF-Tu/Phe-tRNA/GMPPNP ternary complex²² to overlapping binding sites on the ribosome^{3–5,23}. A recent cryo-EM study of EF-G binding, in the presence of an antibiotic thiostrepton²⁴, does not show the ratchet-like movement, presumably because thiostrepton, by binding to the GTPase-associated centre²⁵, prevents EF-G from binding in a way required to induce the dramatic conformational change of the ribosome. Furthermore, our findings do not support earlier models^{26,27} according to which translocation occurs simultaneously with the relative movement between the subunits. We also note

that an earlier ratchet model²⁸ deals exclusively with the relative movement of tRNAs and mRNA.

Cryo-EM^{8,18} and X-ray¹⁹ maps of the 70S ribosome show a number of bridges connecting the subunits, which could have either passive (registration, elastic storage of energy) or active (transmission of force, signalling) roles in the control of the subunit movements. The bridges most strongly affected by the inter-subunit rotation, B1a, B1b and B1c, link the 30S subunit head with the central protuberance of the 50S subunit^{18,19}, whose conformation changes dramatically (Fig. 3). Our results, obtained by direct visualization, explain previous experimental findings^{6,7,29,30}, which were taken to indicate the existence of different relative constellations of two ribosomal subunits during translation. □

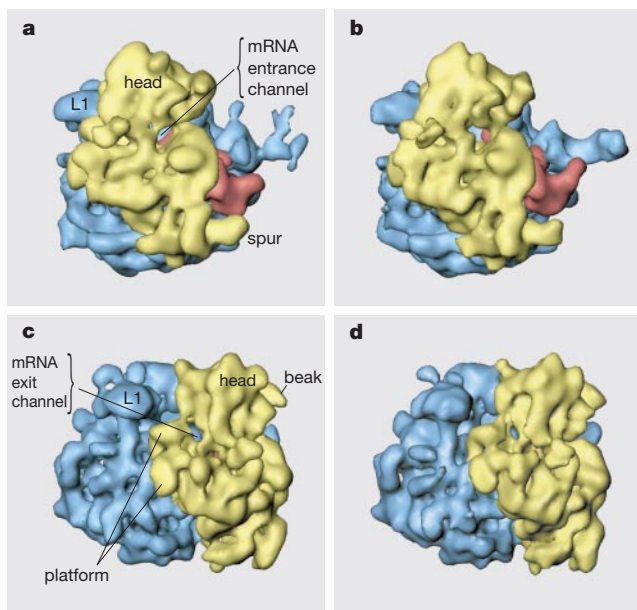


Figure 2 Effect of EF-G-dependent GTP hydrolysis on the dynamics of the mRNA channel. **a, c**, 70S/EF-G/GMPP(CH₂)P complex; and **b, d**, 70S/EF-G/GDP/fusidic acid complex. **a, b**, Maps of 70S ribosome from the channel entrance side, on the solvent side near the

shoulder, in an orientation different to that in Fig. 1 to highlight the channel opening. **c, d**, Maps from the channel exit side, or platform side of the 30S subunit.

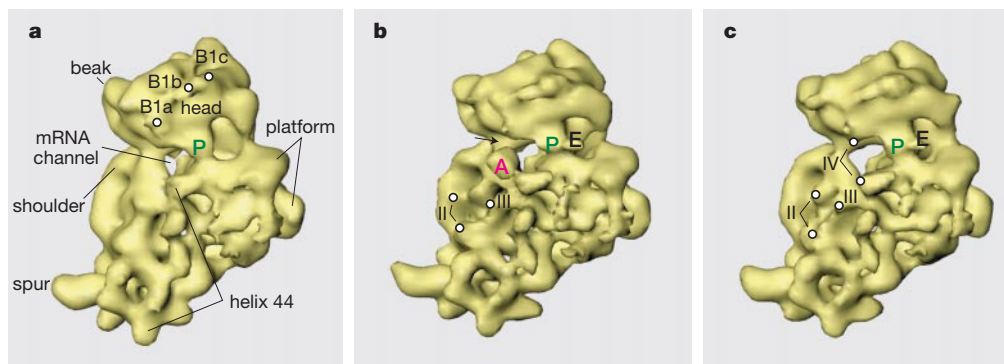


Figure 3 Effect of rotation on the tRNA position before and after GTP hydrolysis. The 30S subunit from various complexes is shown from the subunit interface side, that is, rotated by about 180 degrees around the vertical axis of the view in Fig. 1. Subsequently, the subunit is slightly rotated around the horizontal axis to show the mRNA entrance channel. The 50S subunit, EF-G and portions of tRNAs have been removed computationally. Only portions of the tRNAs making contact with the 30S subunit are retained. **a**, 70S/tMet-tRNA^{Met} complex, showing the P site filled. Upper (decoding site) and lower ends of helix 44 of 16S RNA^{18–20} are marked. **b**, 70S/(tRNA)₂/EF-G/GMPP(CH₂)P complex, showing both A and P sites filled, even though the 30S subunit has

already rotated with respect to the 50S subunit. In this panel, mRNA entrance channel is partially obscured by the density of the A-site tRNA. **c**, 70S/(tRNA)₂/EF-G/GDP/fusidic acid complex, showing the A site vacated but the P site occupied. Landmarks: B1a, B1b and B1c, areas of the 30S subunit head involved in formation of bridges^{18,19} with the central protuberance of the 50S subunit in the control; A, portion of the A-site tRNA contacting the subunit; P, portion of the P-site tRNA contacting the head and platform of the subunit; E, anticodon portion of the E-site tRNA. II, III and IV indicate the areas of the 30S subunit that make contact with EF-G domains II and III in the GTP state (**b**), and II, III, and IV in the GDP state^{3–5} (**c**).

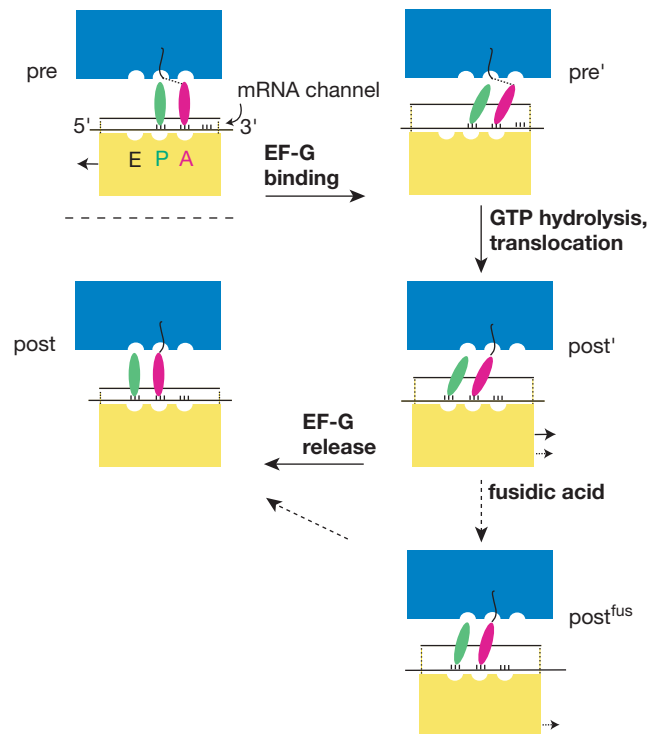


Figure 4 Diagram showing three steps of the translocation process. 'pre' state: tRNAs are in A (magenta) and P (green) sites on both 30S (yellow) and 50S (blue) subunits; the peptide bond has already been formed and the elongated peptidyl moiety is now on the A-site tRNA; the subunits are in the original orientations; both entrance and exit channels for mRNA, indicated by its 3' and 5' ends, are narrow. 'pre'' state, following EF-G binding: 30S subunit has rotated with respect to the 50S subunit, without changing the positions of the tRNAs relative to either subunit; mRNA channel has widened enabling mRNA advance in the next step. 'post' state, following GTP hydrolysis (transient state, not

experimentally observed): tRNA and associated mRNA have advanced from A and P sites to P and E sites (which might involve transitional, A/P and P/E hybrid states²⁹); relative subunit orientations are unchanged. 'post^{fus}' state, assumed when fusidic acid is present: small subunit has partially rotated back, complete reversal is blocked by the presence of EF-G/GDP which is not released. 'post' state, assumed after EF-G is released under normal conditions: small subunit is fully rotated back into original orientation; tRNAs are in P and E sites; mRNA channel is tight again, securing mRNA in its new position.

Methods

Five separate 3D cryo-EM reconstructions of ribosomal complexes^{2,3} with resolutions varying from 15 Å to 18 Å have been analysed. Preparation of complexes, electron microscopy and image processing were described in detail previously²⁻⁴ (70S/fMet-tRNA^{Met} complex²; 70S/EF-G/GMPP(CH₂)P complex³; 70S/(tRNA)₃/EF-G/GMPP(CH₂)P complex³; 70S/EF-G/GDP/fusidic acid complex^{3,4}; and 70S/(tRNA)₂/EF-G/GDP/fusidic acid complex³). For the quantitative comparison, all density maps were individually filtered such that their radial Fourier amplitude profile matched the profile obtained by X-ray solution scattering of the 70S ribosome¹⁸. Finally, all maps were low-pass filtered to a common resolution of 18 Å, as determined by the Fourier shell correlation using a threshold cut-off of 0.5 (ref. 2). The final common resolution of these maps would be 13 Å if measured and filtered by the less conservative 3-σ criterion used by others in the field^{23,24} (see ref. 2 for a discussion of methods of calculating resolution of cryo-EM maps). Radii of gyration of all three-dimensional maps were calculated using the SPIDER software³¹. An animated sequence alternating between the control and the 70S/EF-G/GMPP(CH₂)P complex while rotating around a vertical axis was computed using the 3D modelling/animation software (Lightwave 3D, version 5.5, NewTek Inc., San Antonio, Texas). The animation is available as Supplementary Information.

Received 30 March; accepted 15 June 2000.

- Wilson, K. S. & Noller, H. F. Molecular movement inside the translational engine. *Cell* **92**, 337–349 (1998).
- Malhotra, A. *et al.* *Escherichia coli* 70 S ribosome at 15 Å resolution by cryo-electron microscopy: localization of fMet-tRNA^{Met} and fitting of L1 protein. *J. Mol. Biol.* **280**, 103–116 (1998).
- Agrawal, R. K., Heagle, A. B., Penczek, P., Grassucci, R. A. & Frank, J. EF-G-dependent GTP hydrolysis induces translocation accompanied by large conformational changes in the 70S ribosome. *Nature Struct. Biol.* **6**, 643–647 (1999).
- Agrawal, R. K., Penczek, P., Grassucci, R. A. & Frank, J. Visualization of elongation factor G on the *Escherichia coli* 70S ribosome: the mechanism of translocation. *Proc. Natl Acad. Sci. USA* **95**, 6134–6138 (1998).
- Agrawal, R. K., Heagle, A. B. & Frank, J. in *The Ribosome: Structure, Function, Antibiotics, and Cellular Interactions* (eds Garrett, R. A. *et al.*) 53–62 (ASM, Washington DC, 2000).
- Spirin, A. S., Baranov, V. I., Polubov, G. S., Serdyuk, I. N. & May, R. P. Translocation makes the ribosome less compact. *J. Mol. Biol.* **194**, 119–128 (1987).

- Serdyuk, I. *et al.* Structural dynamics of translating ribosomes. *Biochimie* **74**, 299–306 (1992).
- Frank, J. *et al.* A model of protein synthesis based on cryo-electron microscopy of the *E. coli* ribosome. *Nature* **376**, 441–444 (1995).
- Agrawal, R. K. *et al.* Direct visualization of A-, P-, and E-site transfer RNAs in the *Escherichia coli* ribosome. *Science* **271**, 1000–1002 (1996).
- Mueller, F., Stark, H., van Heel, M., Rinke-Appel, J. & Brimacombe, R. A new model for the three-dimensional folding of *Escherichia coli* 16 S ribosomal RNA. III. The topography of the functional centre. *J. Mol. Biol.* **271**, 566–587 (1997).
- Frank, J. *et al.* in *The Ribosome: Structure, Function, Antibiotics, and Cellular Interactions* (eds Garrett, R. A. *et al.*) 45–51 (ASM, Washington DC, 2000).
- Lata, K. R. *et al.* Three-dimensional reconstruction of the *Escherichia coli* 30 S ribosomal subunit in ice. *J. Mol. Biol.* **262**, 43–52 (1996).
- Agrawal, R. K., Lata, K. R. & Frank, J. Conformational variability in *Escherichia coli* 70S ribosome as revealed by 3D cryo-electron microscopy. *Int. J. Biochem. Cell Biol.* **31**, 243–254 (1999).
- Powers, T. & Noller, H. F. The 530 loop of 16S rRNA: a signal to EF-Tu? *Trends Genet.* **10**, 27–31 (1994).
- Newcomb, L. F. & Noller, H. F. Directed hydroxyl radical probing of 16S rRNA in the ribosome: spatial proximity of RNA elements of the 3' and 5' domains. *RNA* **5**, 849–855 (1999).
- Rodnina, M. V., Savelsbergh, A., Katunin, V. I. & Wintermeyer, W. Hydrolysis of GTP by elongation factor G drives tRNA movement on the ribosome. *Nature* **385**, 37–41 (1997).
- Agrawal, R. K. *et al.* Visualization of tRNA movements on the *E. coli* ribosome during the elongation cycle. *J. Cell. Biol.* (in the press).
- Gabashvili, I. S. *et al.* Solution structure of the *E. coli* 70S ribosome at 11.5 Å resolution. *Cell* **100**, 537–549 (2000).
- Cate, J. H., Yusupov, M. M., Yusupova, G. Z., Earnest, T. N. & Noller, H. F. X-ray crystal structures of 70S ribosome functional complexes. *Science* **285**, 2095–2104 (1999).
- Clemons, W. M. *et al.* Structure of a bacterial 30S ribosomal subunit at 5.5 Å resolution. *Nature* **400**, 833–840 (1999).
- Stark, H. *et al.* Arrangement of tRNAs in pre- and posttranslational ribosomes revealed by electron cryomicroscopy. *Cell* **88**, 19–28 (1997).
- Nissen, P. *et al.* Crystal structure of the ternary complex of Phe-tRNA^{Phe}, EF-Tu, and a GTP analog. *Science* **270**, 1464–1472 (1995).
- Stark, H. *et al.* Visualization of elongation factor Tu on the *Escherichia coli* ribosome. *Nature* **389**, 403–406 (1997).
- Stark, H., Rodnina, M. V., Wieden, H. -J., van Heel, M. & Wintermeyer, W. Large scale movement of elongation factor G and extensive conformational change of the ribosome during translocation. *Cell* **100**, 301–309 (2000).

25. Porse, B. T., Levieu, I., Mankin, A. S. & Garrett, R. A. The antibiotic thiostrepton inhibits a functional transition within protein L11 at the ribosomal GTPase center. *J. Mol. Biol.* **276**, 391–404 (1998).
26. Bretscher, M. S. Translocation in protein synthesis: A hybrid structure model. *Nature* **218**, 675–677 (1968).
27. Spirin, A. S. A model of the functioning ribosome: locking and unlocking of the ribosome subparticles. *Cold Spring Harbor Symp. Quant. Biol.* **34**, 197–207 (1969).
28. Woese, C. Molecular mechanics of translation: a reciprocating ratchet mechanism. *Nature* **226**, 817–820 (1970).
29. Moazed, D. & Noller, H. F. Intermediate states in the movement of transfer RNA in the ribosome. *Nature* **342**, 142–148 (1989).
30. Agrawal, R. K. & Burma, D. P. Sites of ribosomal RNAs involved in the subunit association of tight and loose couple ribosomes. *J. Biol. Chem.* **271**, 21285–21291 (1996).
31. Frank, J. *et al.* SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J. Struct. Biol.* **116**, 190–199 (1996).

Supplementary information is available on Nature's World-Wide Web site (<http://nature.com>).

Acknowledgements

This work was supported by grants from the National Institutes of Health. We thank A. Heagle for preparing the illustrations and animated sequence, and I. Gabashvili and P. Penczek for help with image processing.

Correspondence and requests for materials should be addressed to J.F.

Mimicry of ice structure by surface hydroxyls and water of a β -helix antifreeze protein

Yih-Cherng Liou, Ante Tocilj, Peter L. Davies & Zongchao Jia

Department of Biochemistry, Queen's University, Kingston, Ontario K7L 3N6, Canada

Insect antifreeze proteins (AFP) are much more effective than fish AFPs at depressing solution freezing points by ice-growth inhibition^{1,2}. AFP from the beetle *Tenebrio molitor* is a small protein (8.4 kDa) composed of tandem 12-residue repeats³ (TCTxSxxCxxAx). Here we report its 1.4-Å resolution crystal structure, showing that this repetitive sequence translates into an exceptionally regular β -helix. Not only are the 12-amino-acid loops almost identical in the backbone, but also the conserved side chains are positioned in essentially identical orientations, making this AFP perhaps the most regular protein structure yet observed. The protein has almost no hydrophobic core but is stabilized by numerous disulphide and hydrogen bonds. On the conserved side of the protein, threonine-cysteine-threonine motifs are arrayed to form a flat β -sheet, the putative ice-binding surface. The threonine side chains have exactly the same rotameric conformation and the spacing between OH groups is a near-perfect match to the ice lattice. Together with tightly bound co-planar external water, three ranks of oxygen atoms form a two-dimensional array, mimicking an ice section.

AFP from the beetle *Tenebrio molitor* (TmAFP) is shaped like a slightly flattened cylinder, 32 Å in length, and 6.5×14 Å in the pseudo-rectangular cross section (Fig. 1). The underlying architecture of this 8.4-kDa, threonine- and cysteine-rich protein is that of an extremely regular parallel β -helix formed from seven 12-amino-acid turns or loops. This structure is designed to present a flat, rigid β -sheet along one side of the molecule, which we propose is the ice-binding site. With the exception of the amino-terminal turn, each helix turn contributes a short β -strand to the sheet, typically TCT, with the threonine residues projecting outwards in two precisely aligned parallel arrays (Fig. 1b). The right-handed β -helix is

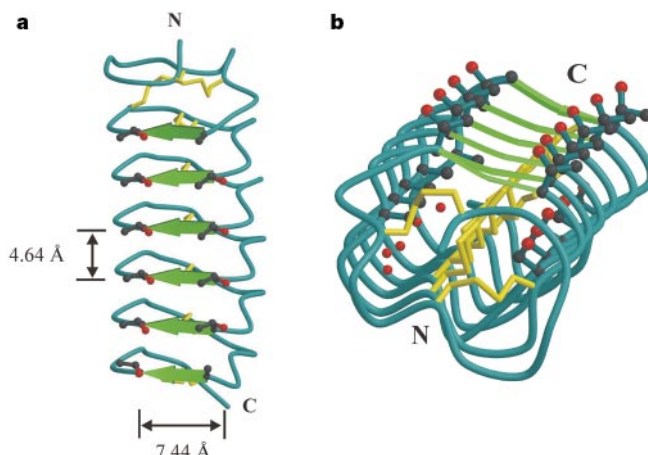


Figure 1 Ribbon illustrations of TmAFP. **a**, Side view of the TmAFP β -helix with the β -sheets (TCT sequences) indicated by green arrows and the disulphide bonds in yellow. Threonine side chains on the β -sheet surface are shown with oxygen atoms in red. **b**, End-on view of the β -helix with the N terminus proximal, showing the alignment of conserved threonine, cysteine, serine and alanine side chains and internal water. We note that the shorter dimension of the pseudo-rectangular cross-section in TmAFP is longer than the disulphide linkage, but the flatness of the β -sheet is maintained by the opposite side being pulled inwards. The flatness of the β -sheet is probably due to the shortness of the β -strands, the disulphide bonds and the presence of favourable van der Waals interactions between stacked threonine side chains. The N and C termini are shown. All diagrams were generated using MOLSCRIPT²⁵ unless otherwise stated.

stabilized by an extensive network of hydrogen bonds parallel to the helix axis, linking peptide CO and NH groups of adjacent turns. The extreme similarity of each helix turn enables inter-loop hydrogen bonds to form at many points around the helix, not just in the β -strand region. Indeed, the root mean square deviation (r.m.s.d.) between six of the seven helical turns (excluding the N-terminal one) is only 0.48 ± 0.02 Å, demonstrating the nearly identical backbone structures. TmAFP is additionally stabilized or constrained by disulphide bonds, which run like rungs of a ladder between opposite sides of the helix (Fig. 1b), making it one of the most extensively disulphide-bonded proteins. As previously predicted⁴, all 16 cysteines of recombinant TmAFP are paired to form disulphide bridges, C3-C12, C9-C19, C16-C22, C28-C34, C40-C46, C52-C58, C64-C70, and C76-C82, equivalent to those seen in the native AFP from the bark beetle, *Dendroides canadensis*^{5,6}. Six of the eight disulphide bonds are spaced six residues apart and are in near-perfect alignment with each other (Fig. 1). In the N-terminal region the other two disulphides (C3-C12, C9-C19) do not fit this pattern, but remarkably this region folds like the rest of the molecule to continue the repetitive β -helix structure with 12 residues per loop (Fig. 1b).

The tight β -helical turns that form the building blocks of TmAFP correspond to the repeated-in-tandem consensus sequence³, TCTxSxxCxxAx. Its identification as a structural unit with an internal disulphide bond is consistent with the observation that different isoforms of TmAFP contain insertions or deletions of this 12-amino-acid repeat^{3,4}. The TCT (or ACT) residues of the repeat represent the β -strand region of the turn, in which the inward-pointing cysteine is disulphide bonded to the other conserved cysteine in order to form a highly constrained seven-residue closed circular loop within the turn (Fig. 2). The extraordinary tightness of the 12-amino-acid turns is also facilitated by intra-loop hydrogen bond connections between backbone CO and NH groups. At all four corners of the pseudo-rectangular cross-section, regular secondary turn structures are formed (γ -, γ -, γ - and β -turns).

Although a number of parallel β -helix structures have been

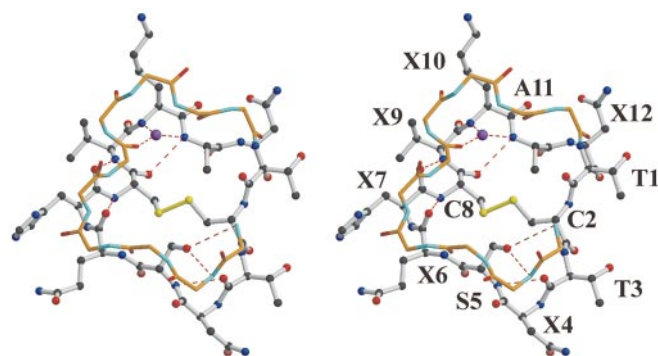


Figure 2 Substructure of a 12-amino-repeat from TmAFP (white) forming one complete loop of the β -helix. Conserved amino acids are identified by their one-letter code and variable residues by X. Numbers indicate their position in the 12-amino-acid repeat. Backbone intra-loop hydrogen bonds (red dashed lines) denote γ - and β -turns. A portion of the preceding repeat is shown (orange) to illustrate the inter-loop hydrogen-bond connections formed by the internal water and serine side chains. Nitrogen atoms are shown in blue and oxygen atoms in red.

reported in bacteria and fungi, including the right-handed β -helices of pectate lyase⁷, pectin lyase⁸, rhamnogalacturonase A⁹ and tail-spike endorhamnosidase of phage P22¹⁰, and the left-handed β -helix domain of *N*-acetylglucosamine acyltransferase¹¹, none of these structures is as compressed or regular as TmAFP. These enzymes contain a larger number of residues per turn of β -helix (typically from 22 to 27)¹², and always have a hydrophobic core. In contrast, with only 12 amino acids per turn, TmAFP is the smallest β -helix and has a very narrow bore, which is constricted and further bisected by disulphide bonds to form two channels, leaving no room for a hydrophobic core. Thus, the few hydrophobic residues in TmAFP—V26, V35, F59 and Y71 (the site of iodination)—all have their side chains projecting outwards to the solvent.

In TmAFP, there is room only for the relatively small side chains of the conserved serine and alanine to project into the core, on either side of the bisecting disulphide bridge. The serine side chains are perpendicular to the disulphide bridge (Figs 1b and 2), with their hydroxyls arrayed in exactly the same orientation ($\chi_1 = -64.0 \pm 1.0^\circ$) at an average distance of 4.73 ± 0.05 Å apart. These serine hydroxyl groups tilt towards the N terminus and form inter-turn hydrogen bonds to two backbone NH groups on the N-terminal neighbouring turn. In the channel opposite serine, all alanine side chains are perpendicular to the disulphide bond, with their methyl groups aligned and separated by an average distance of 5.39 ± 0.05 Å. Within the alanine-containing channel there are also five bound waters regularly spaced between the six carboxy-terminal loops. These bound waters play a similar stabilizing role to the OH group of the internal serine residues, each forming hydrogen bonds to three residues. With an average *B*-factor of 14.3 Å² (as compared to the average *B* of 16.6 Å² for protein atoms and the average *B* of 35.4 Å² for all other waters), these internal water molecules are very tightly bound and can be regarded as an intrinsic part of the protein.

The β -helix structure of TmAFP is very different from any of the deduced or predicted fish AFP structures¹³. But interestingly, it is similar to the structure of another insect AFP determined by NMR². This equally hyperactive 9-kDa spruce budworm AFP (sbwAFP) also forms a β -helix, probably as a result of convergent evolution. The main structural feature that these two insect AFPs have in common is the regular array of TXT sequences on one side of the molecule. In sbwAFP, this has been identified as the ice-binding site by site-directed mutagenesis². In TmAFP, the six short β -strands (Fig. 1b) form a remarkably flat β -sheet that lacks the typical twist

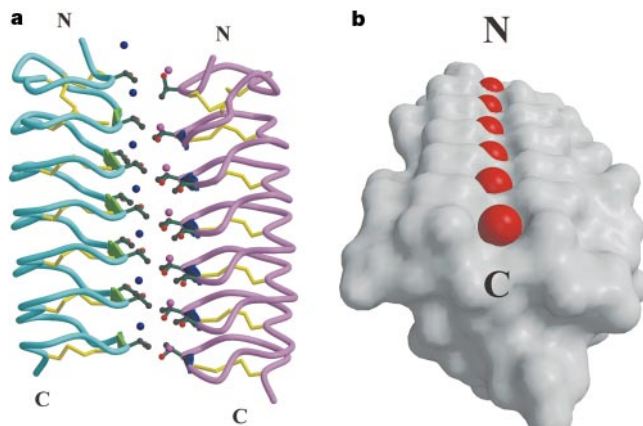


Figure 3 Dimer of TmAFP and organization of external water. **a**, A TmAFP pair dimerized through hydrogen bonding to two ranks of ordered waters. The first TmAFP is shown in blue; the second in pink. The oxygen atoms of the waters are shown as blue and magenta spheres. **b**, A surface presentation (GRASP²⁶) of TmAFP with the closest rank of water molecules (red spheres) bound, illustrating their regularity and the flatness of the resulting surface when one rank of the water molecules is in place.

seen in other β -sheet proteins. Our high-resolution TmAFP structure shows that the threonine residues are precisely positioned and fully exposed to the solvent, with their side chains in near-perfect alignment ($\chi_1 = -58.2 \pm 0.8^\circ$). The average distance between hydroxyls within the TCT motif is 7.44 ± 0.05 Å. At right angles to this dimension, the average distance between equivalent hydroxyls of TCT motifs in adjacent loops is 4.64 ± 0.04 Å. The two-dimensional array of threonine side chains makes a remarkably good match to the repeated spacing between oxygen atoms in the ice lattice on the primary prism plane (7.35 Å and 4.52 Å), and a reasonable match to the basal plane (7.83 Å and 4.52 Å). This lattice matching is highly suggestive of its ice-binding function.

Although TmAFP is monomeric in solution¹⁴, the protein crystallized as a dimer with twofold symmetry. The two monomers are essentially identical with an r.m.s.d. of only 0.33 Å (including all regular internal and external waters). The dimerization of TmAFP occurs along the surface of the parallel β -sheets, with both monomers in the same (that is, parallel) orientation (Fig. 3a). The two monomers are at least 8 Å apart and do not directly interact with each other. Rather, contact between the two monomers is mediated by two highly ordered ranks of water molecules that form an extensive hydrogen-bonding network with the threonine residues on one side of the TCT motifs (not shown). The remarkable regularity of the water molecules (Fig. 3b) is reflected by their near-identical positions in both monomers (0.46 Å r.m.s.d.). Each water molecule forms two hydrogen bonds to the closer monomer and one hydrogen bond to the more distant monomer. These external water molecules are tightly bound and highly ordered, with an average *B* factor of only 16.5 Å².

Because of the positional regularity, the distance between two adjacent waters is 4.64 ± 0.20 Å, the same distance as between threonine residues (4.64 ± 0.23 Å). A good two-dimensional match to the ice lattice, including all three ranks of oxygen atoms (two from threonine and one from water), can be achieved with <0.5 Å offset or deviation, implying that the ordered water molecules could act as part of an ice surface and directly participate in the AFP–ice interaction (Fig. 4). Importantly, this match requires no adjustment of any threonine side chain and methyl groups of threonine, and side chains of adjacent residues do not present steric interference (Fig. 4a). Hence the regular array formed by three ranks of oxygen atoms can be seen as a small piece of one-layer-thick ice to be incorporated into a large ice lattice, an idea first proposed by Knight *et al.*¹⁵. Regardless of whether the bound water

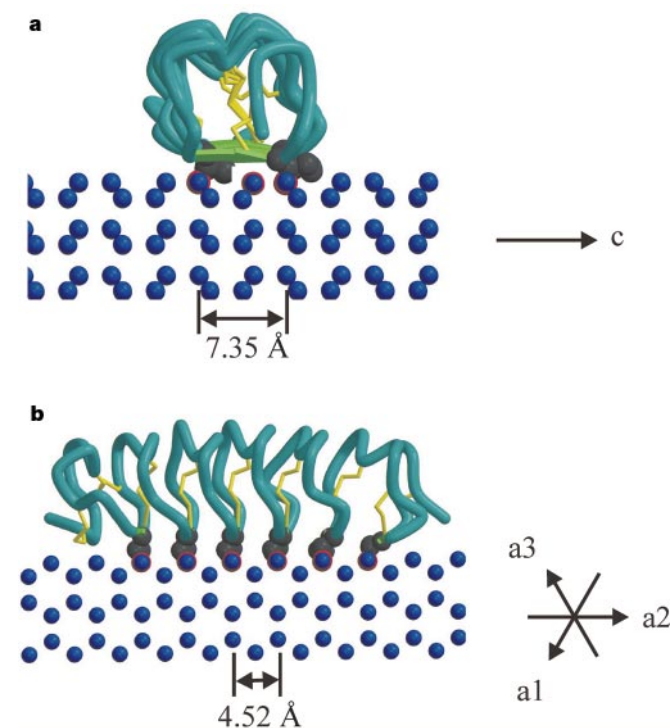


Figure 4 Lattice matching/occupation model for TmAFP binding to ice. **a**, End-on view of the β -helix with the C terminus proximal, showing the occupation of ice oxygen atoms (blue) on the prism plane of ice by three ranks of oxygen atoms from threonine and bound water (red perimeter). We note the absence of steric clashes with the threonine methyl groups. **b**, Side view of the β -helix with the C terminus to the right, showing the occupation of ice oxygen atoms on the prism plane of ice by threonine from six loops. Ice c-axis and a-axis directions are indicated by the arrows.

molecules are considered to be part of the protein or part of the ice, their arrangement (along with the threonine residues) to match the ice lattice represents the first direct observation, to our knowledge, of the structure of an AFP–ice interface.

Dimerization of TmAFP may have fortuitously stabilized the bound water to give us this first glimpse and approximation of the elusive AFP–ice structure. Certainly, bound water matching the ice lattice has not been seen before in the X-ray structures of the two fish AFPs, types I and III (refs 16, 17). In comparison to the single α -helix structure of fish type I AFP, which contains one array of regularly spaced threonine residues, TmAFP possesses multiple ranks of regularly spaced threonine residues and bound water molecules. If it is remarkable that lattice-matching regularity can occur in one dimension, as in type I AFP, then it is even more surprising for a protein to achieve the extreme regularity in two dimensions that is seen in TmAFP. A highly constrained β -helix is perhaps an ideal platform for the design of two-dimensional regularity. □

Methods

We have previously reported the preparation and crystallization of iodinated recombinant TmAFP¹⁸. The structure determined is that of the Y71-iodinated TmAFP. X-ray data including anomalous reflections were collected to 1.95-Å resolution using a rotating anode X-ray generator with an imaging plate. TmAFP crystals belonged to a $P6_5$ space group, with unit cell dimensions $a = b = 73.1$ and $c = 53.1$ Å. There are two molecules per asymmetric unit and there is a twofold non-crystallographic symmetry. Using direct methods implemented¹⁹ in SHELX-97, the positions of four iodine atoms were determined and the correct space group $P6_5$ was assigned. The iodine sites were refined, the phases were subsequently obtained using the method of single anomalous scattering and a solvent-flattened map (51% solvent) was generated using the program SHARP²⁰. In addition, figures of merit generated from two opposite enantiomers were examined to

determine handedness²¹. To improve the map, non-crystallographic averaging and density modification were performed. Model building was carried out using the program O²², and final refinement was carried out using the program CNS²³ with the 1.4-Å data collected at the X8C beamline of Brookhaven National Laboratory. An individual isotropic B -factor model with non-crystallographic symmetry restraint was used in the initial refinement. The R factor of the initial model with 154 water molecules was improved from 36.0% to 23.54% ($R_{\text{free}} = 25.24\%$). A further refinement was then carried out using SHELX-97¹⁹. After the refinement of anisotropic B -factors and occupancies of iodine atoms were introduced, the R -factor dropped to 16.99% ($R_{\text{free}} = 20.73\%$). The R -factor of the final model, containing 164 amino acid residues of the dimer (N-terminal methionine and C-terminal glycine-histidine were invisible) and 229 water molecules, is 16.05% ($R_{\text{free}} = 19.96\%$). The structure geometry is excellent with no Ramachandran violations²⁴. Calculations of r.m.s.d. included all main-chain atoms, where appropriate.

Received 9 March; accepted 15 May 2000.

1. Tyshenko, M. G., Doucet, D., Davies, P. L. & Walker, V. K. The antifreeze potential of the spruce budworm thermal hysteresis protein. *Nature Biotechnol.* **15**, 887–890 (1997).
2. Graether, S. P. *et al.* β -Helix structure and ice-binding properties of a hyperactive antifreeze protein from an insect. *Nature* **406**, 325–328 (2000).
3. Graham, L. A., Liou, Y. -C., Walker, V. K. & Davies, P. L. Hyperactive antifreeze protein from beetles. *Nature* **388**, 727–728 (1997).
4. Liou, Y. -C., Thibault, P., Davies, P. L., Walker, V. K. & Graham, L. A. A complex family of highly heterogeneous and internally repetitive hyperactive antifreeze proteins from the beetle *Tenebrio molitor*. *Biochemistry* **38**, 11415–11424 (1999).
5. Li, N., Chibber, B. A. K., Castellino, F. J. & Duman, J. G. Mapping of disulfide bridges in antifreeze proteins from overwintering larvae of the beetle *Dendroides canadensis*. *Biochemistry* **37**, 6343–6350 (1998).
6. Duman, J. G. *et al.* Molecular characterization and sequencing of antifreeze proteins from larvae of the beetle *Dendroides canadensis*. *J. Comp. Physiol. B* **168**, 225–232 (1998).
7. Yoder, M. D., Keen, N. T. & Jurnak, F. New domain motif: structure of pectate lyase C, a secreted plant virulence factor. *Science* **260**, 1503–1507 (1993).
8. Mayans, O. *et al.* Two crystal structures of pectin lyase A from *Aspergillus* reveal a pH driven conformational change and striking divergence in the substrate-binding clefts of pectin and pectate lyases. *Structure* **5**, 677–689 (1997).
9. Petersen, T. N., Kauppinen, S. & Larsen, S. The crystal structure of rhamnogalacturonase A from *Aspergillus aculeatus*: a right-handed parallel beta helix. *Structure* **5**, 533–544 (1997).
10. Steinbacher, S. *et al.* Crystal structure of P22 tailspike protein: interdigitated subunits in a thermostable trimer. *Science* **265**, 383–386 (1994).
11. Raetz, C. R. & Roderick, S. L. A left-handed parallel beta helix in the structure of UDP-N-acetylglucosamine acyltransferase. *Science* **270**, 997–1000 (1995).
12. Jenkins, J., Mayans, O. & Pickersgill, R. Structure and evolution of parallel beta-helix proteins. *J. Struct. Biol.* **122**, 236–246 (1998).
13. Davies, P. L. & Sykes, B. D. Antifreeze proteins. *Curr. Opin. Struct. Biol.* **7**, 828–834 (1997).
14. Liou, Y. -C. *et al.* Folding and structural characterization of highly disulfide-bonded beetle antifreeze protein produced in bacteria. *Protein Expr. Purif.* **19**, 148–157 (2000).
15. Knight, C. A., Driggers, E. & DeVries, A. L. Adsorption to ice of fish antifreeze glycopeptides 7 and 8. *Biophys. J.* **64**, 252–259 (1993).
16. Sicheri, F. & Yang, D. S. C. Ice-binding structure and mechanism of an antifreeze protein from winter flounder. *Nature* **375**, 427–431 (1995).
17. Jia, Z., DeLuca, C. L., Chao, H. & Davies, P. L. Structural basis for the binding of a globular antifreeze protein to ice. *Nature* **384**, 285–288 (1996).
18. Liou, Y. -C., Davies, P. L. & Jia, Z. Crystallization and preliminary X-ray analysis of insect antifreeze protein from the beetle, *Tenebrio molitor*. *Acta Cryst. D* **56**, 354–356 (2000).
19. Sheldrick, G. M. & Schneider, T. R. in *Methods in Enzymology* Vol. 276 (eds Carter, C. W. & Sweet, R. M.) 319–343 (Academic Press, New York, 1997).
20. de La Fortelle, E. & Brice, G. in *Methods in Enzymology* Vol. 276 (eds Carter, C. W. & Sweet, R. M.) 472–494 (Academic Press, New York, 1997).
21. Chen, L. Q. *et al.* Crystal structure of a bovine neurophysin II dipeptide complex at 2.8 Å determined from the single-wavelength anomalous scattering signal of an incorporated iodine atom. *Proc. Natl Acad. Sci. USA* **88**, 4240–4244 (1991).
22. Jones, T. A., Zou, J. -Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst. A* **47**, 110–119 (1991).
23. Brünger, A. T. *et al.* Crystallography & NMR system (CNS): A new software suite for macromolecular structure determination. *Acta Cryst. D* **54**, 905–921 (1998).
24. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).
25. Kraulis, P. J. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
26. Nicholls, A., Sharp, K. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281–296 (1991).

Acknowledgements

We would like to thank Q. Ye, M. Kuiper and S. Gauthier for excellent technical assistance, and staff at the X8C beamline of Brookhaven National Laboratory for help with synchrotron data collection. This work was supported by grants from MRC of Canada to Z. J and P. L. D; Z. J. is an MRC Scholar. P.L.D. is a Killam Research Fellow.

Correspondence and requests for materials should be addressed to P.L.D. (e-mail: daviessp@post.queensu.ca). The coordinates have been deposited in the Protein Data Bank under accession code 1EZG.

β -Helix structure and ice-binding properties of a hyperactive antifreeze protein from an insect

Steffen P. Graether^{*†}, Michael J. Kuiper[‡], Stéphane M. Gagné[§], Virginia K. Walker[‡], Zongchao Jia^{*}, Brian D. Sykes[§] & Peter L. Davies^{*‡}

^{*} Department of Biochemistry, Queen's University, Kingston, Ontario K7L 3N6, Canada

[‡] Department of Biology, Queen's University, Kingston, Ontario K7L 3N6, Canada

[§] Department of Biochemistry, University of Alberta, Edmonton, Alberta T6G 2H7, Canada

[†] Present address: Department of Biochemistry, University of Alberta, Edmonton, Alberta T6G 2H7, Canada.

Insect antifreeze proteins (AFP) are considerably more active at inhibiting ice crystal growth than AFP from fish or plants. Several insect AFPs, also known as thermal hysteresis proteins, have been cloned^{1–3} and expressed^{1,2}. Their maximum activity is 3–4 times that of fish AFPs¹ and they are 10–100 times more effective at micromolar concentrations. Here we report the solution structure of spruce budworm (*Choristoneura fumiferana*) AFP and characterize its ice-binding properties. The 9-kDa AFP is a β -helix with a triangular cross-section and rectangular sides that form stacked parallel β -sheets; a fold which is distinct from the three known fish AFP structures. The ice-binding side contains 9 of the 14 surface-accessible threonines organized in a regular array of TXT motifs that match the ice lattice on both prism and basal planes. In support of this model, ice crystal morphology and ice-etching experiments are consistent with AFP binding to both of these planes and thus may explain the greater activity of the spruce budworm antifreeze.

Recombinant spruce budworm antifreeze protein (sbwAFP) is as hyperactive as the native protein (Fig. 1), and was used for structure determination. We obtained 1014 nuclear Overhauser effect (NOE) distance and 88 dihedral angle restraints from several nuclear magnetic resonance (NMR) experiments and used them to generate an ensemble of structures (Fig. 2a). The overall fold is that of a left-handed β -helix (Fig. 2b), with one partly structured loop at the amino terminus (loop 1), four complete left-handed loops (loops

2–5) and one right-handed loop at the carboxy terminus (loop 6). The protein fold is stabilized by a hydrophobic core, and by disulphide bonds and parallel β -sheets between the loops. The left-handed loops are 15 residues long, and are spaced 4.5–5.7 Å apart. The right-handed loop is 17 residues long and has two anti-parallel β -sheets that cap the C-terminal end of the protein. In addition, the N, C α and C' atoms of the C-terminal residues 88–90 have a root mean square deviation (r.m.s.d.) of 0.53 Å, which is the same as the r.m.s.d. of the whole protein (Table 1). The cross-section of the protein perpendicular to the axis of the β -helix has a triangular shape. Of the rectangular sides, at least two consist of

Table 1 Structural statistics for the ensemble of 20 refined structures of sbwAFP

	(SA) [*]	(SA) _†
r.m.s.d. from NOE distances (Å) (1014) [¶]	0.012 $\pm$ 0.002	0.014
r.m.s.d. deviations from dihedral angles (degrees) (88)	0.42 $\pm$ 0.07	0.44
r.m.s.d. deviations from carbon chemical shift		
δC^{α} (p.p.m.) (59)	1.43 $\pm$ 0.06	1.45
δC^{β} (p.p.m.) (59)	1.62 $\pm$ 0.06	1.46
r.m.s.d. from idealized covalent geometry		
Bonds (Å)	0.0018 $\pm$ 0.0001	0.0019
Angles (degrees)	0.54 $\pm$ 0.01	0.55
Impropers (degrees)	0.46 $\pm$ 0.02	0.48
Final energies # (kcal mol ⁻¹)		
F_{NOE}	8 $\pm$ 2	9.8
F_{dih}	0.9 $\pm$ 0.2	1.1
F_{repel}	12 $\pm$ 1	12.8
$E_{L-J}^{\ddagger}$	-240 $\pm$ 30	-244
Cartesian coordinate r.m.s.d. (Å) § (SA) versus (SA) _{avg}	N, C α and C'	All heavy atoms
	0.52 $\pm$ 0.09	1.0 $\pm$ 0.1
Ramachandran plot ²⁴	All structures	(SA) _r
Residues in most favoured region	47.9	48.0
Residues in additionally allowed region	50.2	49.3

^{*} Represents the average root mean square deviations (r.m.s.d.) for the ensemble.

[†] (SA)_† is the single member of the ensemble that is closest to the average structure.

[‡] The Lennard-Jones potential was not used in the refinement process.

[§] Residues 4–5, 7–18, 23–69, 71–72, 80–90.

^{||} (SA)_{avg} is the geometric average structure of the ensemble.

[¶] Numbers in brackets represent the number of constraints used in the structure calculation.

[#] Final energies are calculated based on contributions from nuclear Overhauser effect (F_{NOE}), dihedral angles (F_{dih}), van der Waals repulsion (F_{repel}) and Lennard-Jones potential (E_{L-J}).

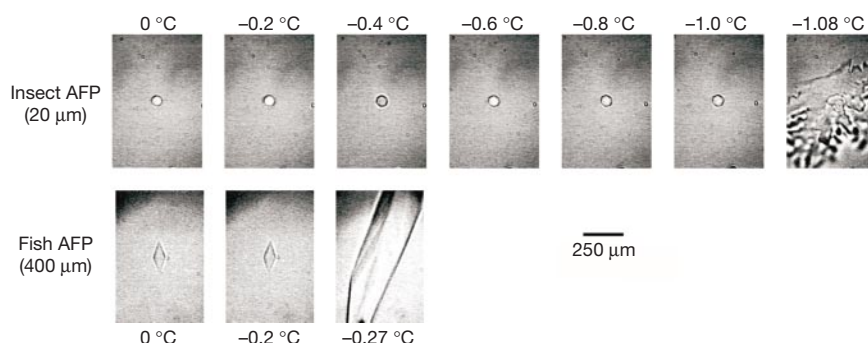


Figure 1 Hyperactivity of insect antifreeze protein compared to a fish antifreeze protein. Ice crystals in the presence of 20 μ m spruce budworm (insect) antifreeze protein (sbwAFP) or 400 μ m winter flounder type I (fish) AFP were observed by video-microscopy during thermal hysteresis measurements at -0.2 °C intervals of cooling, which occurred at a rate of ~0.07 °C per min. There was no observable change in the ice crystals as the temperature was lowered until the non-equilibrium freezing point was exceeded at -1.08 °C for sbwAFP, and at -0.27 °C for type I AFP. Thus, at one-twentieth of the fish

AFP molar concentration, sbwAFP causes about four times the depression of the freezing point (1.08 °C versus 0.27 °C). The hexagonal ice crystal obtained with sbwAFP lies with its *c*-axis perpendicular to the plane of the figure, and the bipyramidal crystal produced by type I AFP has its *c*-axis in the plane of the figure, parallel to its long dimension. Below the non-equilibrium freezing point, uncontrolled growth of the crystal in the presence of sbwAFP occurs laterally along the *a*-axis. In contrast, in the presence of type I fish AFP the crystal grows rapidly along the *c*-axis to form a spicule.

stacked, parallel β -sheets (Fig. 2b). With respect to the third side, 20–30% of the members of the ensemble have β -strand content, suggesting that there is some β -sheeting along this face as well. This third side is threonine-rich due to the clustering and alignment of four TXT motifs (where X is any inward pointing amino acid) at positions 5–7, 21–23, 36–38 and 51–53 (Fig. 2b and c). A fifth motif at 68–70 in the longer loop 5 is imperfect (IXT).

The left-handed β -helices characterized to date consist mostly of 18 residues per loop, but also have a triangular cross-section^{4,5}. The 15-residue loop in sbwAFP is three residues shorter, and corresponds to the removal of one residue from each of the three sides of the triangle. In addition, other β -helices contain secondary structural elements and inserts that extend outward from the helix.

Therefore, sbwAFP and beetle AFP⁶, represent the first isolated β -helix structures determined.

Given the importance of threonines for ice binding in other AFPs^{7–9}, several of these residues in the TXT motifs were mutated to leucine to alter both the polarity and the shape of the protein surface. Mutants T7L, T21L, T38L and T51L each showed an 80–90% loss of activity (Gauthier *et al.*, unpublished results; Fig. 2c). Subsequently, threonines not present in the array of TXT motifs were also mutated. Mutants T48L and T66L, adjacent to the TXT motifs, showed only a moderate decrease in activity (30% and 35%, respectively). Mutation of threonine 86 to leucine, located opposite the TXT side, resulted in no loss of activity. Thus, we propose that the TXT face constitutes the ice-binding site of sbwAFP.

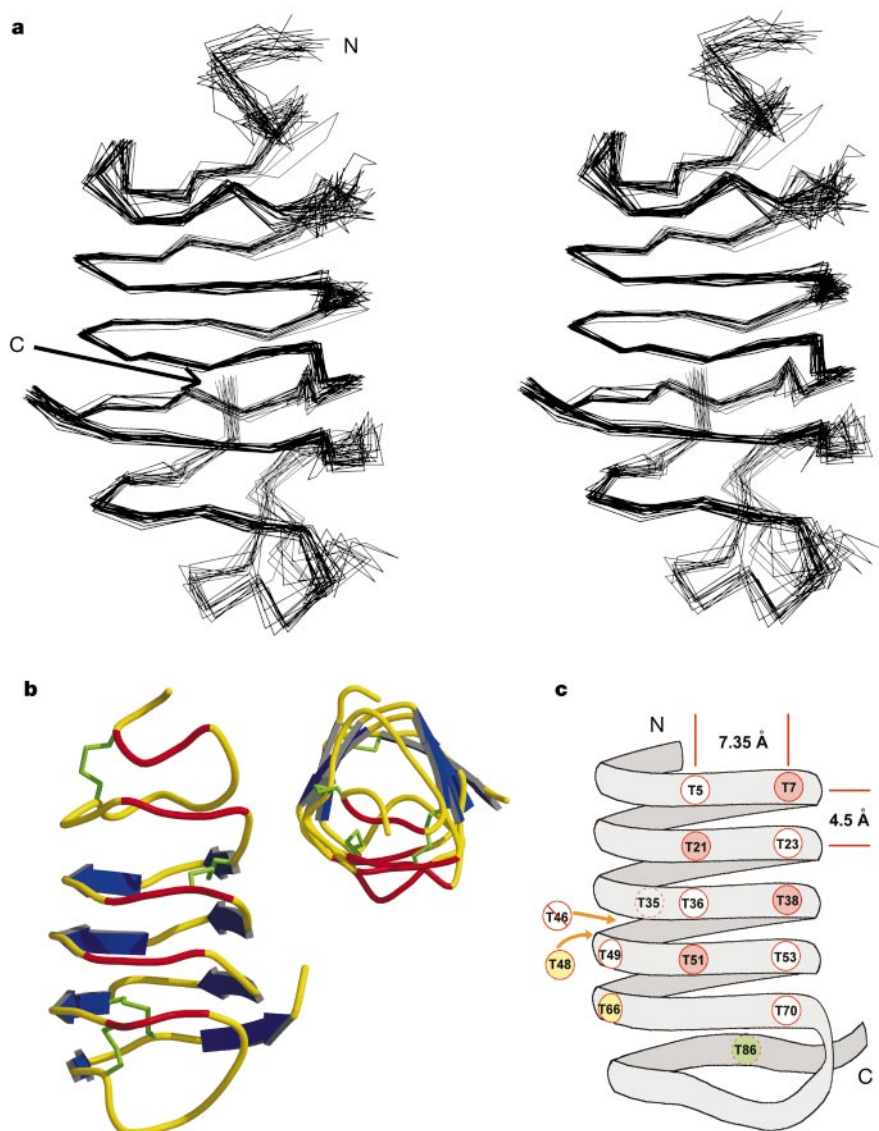


Figure 2 Structure of spruce budworm antifreeze protein. **a**, Stereoview of C α atoms of residues 1–90 from the 20 lowest-energy structures of sbwAFP (of the 40 that converged from the 50 computed). No distance restraint was violated by more than 0.3 Å and no dihedral angle restraint by more than 4°. The structural data statistics are shown in Table 1. Residues in regions 5–8, 19–22, 35–38, 49–53 and 68–70 appear to be exchange-broadened as suggested by preliminary ¹⁵N-relaxation experiments (data not shown), resulting in poorly defined residue positions. **b**, Ribbon representation of the sbwAFP structure closest to the mean. The β -sheets, shown in blue, are located at residues 25–28, 30–33, 40–43, 45–48, 55–58, 62–65, 67–69, 81–83 and 85–89. Disulphide bonds, shown in green, are located between loops 1 and 2, loops 2 and 3, and two

between loops 5 and 6. Coils shown in red indicate the threonine-rich face. The inset shows the view down the β -helix axis. **c**, Schematic representation of sbwAFP showing the relative positions of threonine residues. The putative ice-binding threonines (residues 5, 7, 21, 23, 36, 38, 51, 53 and 70) are aligned in a grid-like arrangement spaced 7.4 Å and 4.5 Å apart. The coloured circles indicate the following percentage antifreeze activities relative to wild-type protein: red, <30%; yellow, <60%; green, 100%; white, not determined. Dotted circles indicate projection of the threonine side chain on the other side of the coil. The line through T46 signifies that the T46L mutation is a folding mutant. The N and C termini are marked.

We performed ice etching¹⁰ and crystal habit studies to determine which ice planes are bound by sbwAFP. Ice hemispheres were grown in the presence of dilute sbwAFP to incorporate the protein into the ice lattice on a binding surface. These binding planes were revealed by partial sublimation of the ice hemisphere at -20°C as 'etches' of AFP accumulation left behind by the receding ice (Fig. 3a). Views of three different hemispheres show sbwAFP binding as distinct etches centred on the six equivalent primary prism planes (p) as well as the basal planes (b) of ice. The etched regions are shaded in the

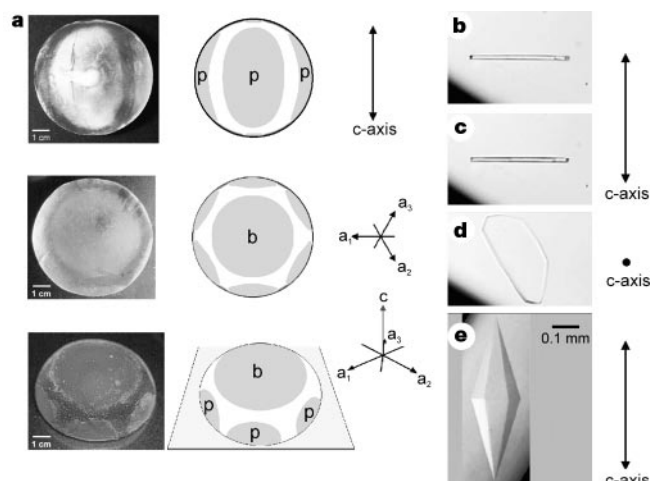


Figure 3 Ice hemispheres and an ice crystal grown in the presence of slow AFP. **a**, The three ice hemispheres were each grown from a single orientated seed crystal in the presence of dilute sbwAFP ($\sim 0.02 \text{ mg ml}^{-1}$). They were subsequently allowed to sublime to reveal bound AFP as a proteinaceous coating (white etch on the ice surface). Alongside are the corresponding interpretive sketches in which the etched surfaces are shaded and identified as a primary prism plane (p) or basal plane (b). The identification as primary prism planes was confirmed by orientation to the facets formed on a drilled cavity after sublimation of the ice hemisphere. The unshaded regions in the sketches represent ice surfaces not bound by AFP, which are most evident in the oblique view in the bottom panel. The relative orientation of the *c*- and *a*-axes of ice is shown on the right. **b**, Single ice crystal grown in the presence of sbwAFP (0.2 mg ml^{-1}) by manipulation of the supercooling temperature and then maintained at -0.2°C . **c**, The same crystal after 24 h. **d**, A tilted view of the same crystal. **e**, Hexagonal bipyramidal morphology produced by 2 mg ml^{-1} of fish type I AFP from the winter flounder. In **b–e**, the direction of the *c*-axis is indicated by arrows, which in **d** is normal to the plane of the figure.

interpretive sketches. The top panel shows three of the six equivalent prism plane etches. The middle panel is a view down the *c*-axis at the basal plane etch, with the tips of the six prism plane etches visible at the periphery. The bottom panel presents an oblique view of an etched hemisphere to show the distinct separation between prism and basal plane etches. Basal plane etching by other AFPs has not been observed before, and we cannot rule out the possibility that protein accumulation on this surface is due to prism plane binding at the risers of steps on the basal plane.

Ice morphology studies at higher sbwAFP concentrations (0.2 mg ml^{-1}) and low supercooling (-0.20°C) also show simultaneous prism-face and basal-plane inhibition (Fig. 3b–d). As crystal habit between the freezing and melting points (thermal hysteresis gap) is the result of growth-inhibited faces, AFPs binding simultaneously to these surfaces typically produce the hexagonal ice crystals seen during thermal hysteresis measurements (Fig. 1, top panels). By teasing some growth out of these crystals before returning them to the thermal hysteresis gap, it is possible to see the growth-inhibited surfaces more clearly. The single crystal viewed perpendicular to the *c*-axis (Fig. 3b) demonstrates that the prism faces are at right angles to the flat basal planes. This crystal showed no growth or shrinkage within the thermal hysteresis gap for at least 24 h (Fig. 3c). The expression and stability of such extensive basal planes is unique to sbwAFP and has not been observed with any of the fish AFPs. Indeed, fish AFPs typically bind to pyramidal or prism planes to form ice crystals with hexagonal bipyramidal or trapezohedral morphology (Fig. 3e)^{10,11}.

Experimental evidence and theories from the fish AFP literature suggest that a specific ice-lattice match is required for AFP binding to ice^{12,13}. The nature of the binding force is not known, but is assumed to involve both hydrogen bonds and non-polar forces^{8,14–17}. Combining NMR structural information and experimental evidence for prism and basal plane binding, we have developed a model to demonstrate the specificity of sbwAFP for these planes (Fig. 4). On the ice-binding site of sbwAFP, the loops are a series of structural repeats, similar to those found in other left-handed β -helix proteins^{4,5}, and bring the TXT motifs into alignment so that the threonines form two parallel arrays. In the model, the spacing between threonines in the neighbouring TXT motifs makes an ideal match to the $4.5 \text{ \AA}$ repeats along the *a*-axis in both the prism and basal planes of ice. At right angles to this repeat, the distance between threonines in the TXT motifs makes an ideal match to the $7.35 \text{ \AA}$ spacing on the prism plane, and a fairly close match to the $7.8 \text{ \AA}$ spacing on the basal plane.

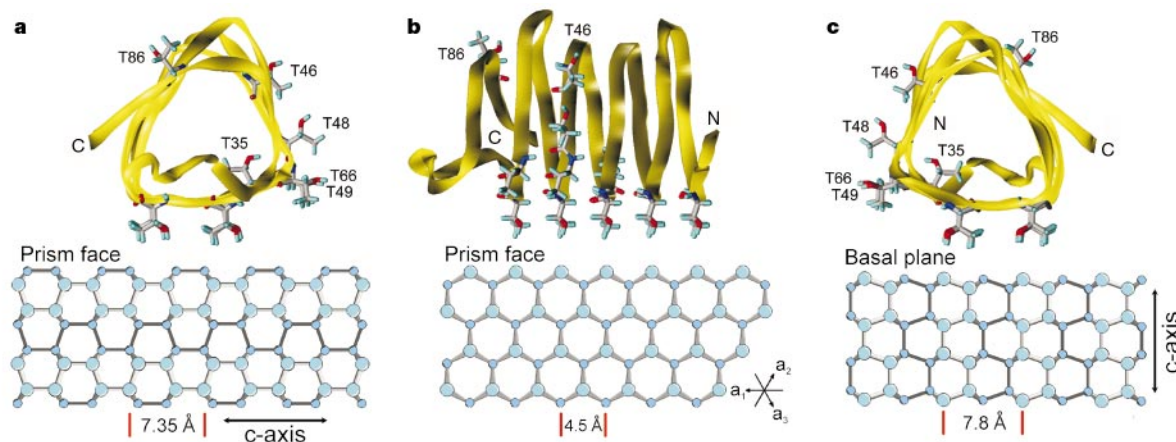


Figure 4 SbwAFP model showing surface complementarity with the prism and basal planes of ice. **a**, A side view of the protein aligned above an ice prism plane. Circles depict the lattice positions of water oxygen atoms in ice. **b**, A view perpendicular to that of **a**. This view down the *c*-axis illustrates the match of adjacent loops to the ice surface. As oxygen

atoms have the same interatomic distance along the *a*-axis in both basal and prism planes, the match to the protein is the same for both planes. **c**, The sbwAFP aligned above the basal plane, depicting the analogous surface match to that of the prism plane.

This simple model demonstrates the geometric feasibility of an ice lattice match between the protein and the primary prism and basal planes of ice without requiring knowledge of the forces involved in the interaction. It is consistent with the mutagenesis, ice-etching and ice-morphology data and shows how the ice-binding site threonines in the aligned TXT motifs could form nine contacts with ice, considerably more than the four contacts modelled between winter flounder AFP and ice¹⁸. The greater number of contacts and the ability to bind both prism and basal planes of ice may explain the greater activity of this insect AFP over fish AFPs.

The sbwAFP and beetle AFP⁶ structures are the first insect AFP structures determined, and the β -helix fold represents a new AFP structural motif. Several different folds have been found for the AFP proteins, suggesting that nature has evolved different structures to form similar ice-binding surfaces. Given the controversy surrounding the nature of the forces involved between fish AFPs and ice^{8,14–17}, the β -helical insect AFPs provide an independent model with which to test binding hypotheses. □

Methods

NMR data collection

Recombinant ¹⁵N- and ¹³C/¹⁵N-labelled sbwAFP was produced using BioExpress media and purified as described¹⁹. The protein was dissolved at a concentration of 18 mg ml⁻¹ in 90% H₂O/10% D₂O or 100% D₂O as necessary, and the pH was adjusted to 5.5 using NaOD or DCl as required. CBCA(CO)NNH, HNCACB, ¹³C/¹⁵N-edited NOESY, ¹⁵N-edited nuclear Overhauser effect enhancement spectroscopy (NOESY), ¹⁵N-edited total correlation spectroscopy (TOCSY), HCCH-TOCSY, HNHA and 2D-NOESY spectra were collected on a Varian 500 or 600 MHz spectrometer with the temperature control set to 30 °C. Data were processed and analysed using NMRPIPE²⁰ and PIPP, and the ensemble of structures was generated in X-PLOR²¹ using distance and dihedral restraints from NOESY and HNHA experiments and empirical restraints from ¹³C-chemical shifts²². Analysis of the ensemble was performed using VADAR²³ and PROCHECK_NMR²⁴.

Ice crystal morphology

Ice hemispheres, initially grown from a single ice crystal, were allowed further growth in the presence of 0.02 mg ml⁻¹ of sbwAFP and permitted to etch as described¹⁰. Ice crystal morphology in the presence of 0.2 mg ml⁻¹ sbwAFP was monitored at -0.2 °C over a period of five days.

Figure preparation

Figures were generated using MOLSCRIPT²⁵, RASTER3D²⁶ and SYBYL²⁷.

Received 22 March; accepted 15 May 2000.

1. Tyshenko, M. G., Doucet, D., Davies, P. L. & Walker, V. K. The antifreeze potential of the spruce budworm thermal hysteresis protein. *Nature Biotechnol.* **15**, 887–890 (1997).
2. Graham, L. A., Liou, Y. C., Walker, V. K. & Davies, P. L. Hyperactive antifreeze protein from beetles. *Nature* **388**, 727–728 (1997).
3. Duman, J. G. *et al.* Molecular characterization and sequencing of antifreeze proteins from larvae of the beetle *Dendroides canadensis*. *J. Comp. Physiol. B* **168**, 225–232 (1998).
4. Beaman, T. W., Sugantino, M. & Roderick, S. L. Structure of the hexapeptide xenobiotic acetyltransferase from *Pseudomonas aeruginosa*. *Biochemistry* **37**, 6689–6696 (1998).
5. Kisker, C., Schindelin, H., Alber, B. E., Ferry, J. G. & Rees, D. C. A left-hand beta-helix revealed by the crystal structure of a carbonic anhydrase from the archaeon *Methanosarcina thermophila*. *Embo J.* **15**, 2323–2330 (1996).

6. Liou, Y. -C., Tocilj, A., Davies, P. L. & Jia, Z. Mimicry of ice structure by surface hydroxyls and water of a β -helix antifreeze protein. *Nature* **406**, 322–324 (2000).
7. Sicheri, F. & Yang, D. S. Ice-binding structure and mechanism of an antifreeze protein from winter flounder. *Nature* **375**, 427–431 (1995).
8. Sönnichsen, F. D., DeLuca, C. I., Davies, P. L. & Sykes, B. D. Refined solution structure of type III antifreeze protein: hydrophobic groups may be involved in the energetics of the protein-ice interaction. *Structure* **4**, 1325–1337 (1996).
9. Jia, Z., DeLuca, C. I., Chao, H. & Davies, P. L. Structural basis for the binding of a globular antifreeze protein to ice. *Nature* **384**, 285–288 (1996).
10. Knight, C. A., Cheng, C. C. & DeVries, A. L. Adsorption of alpha-helical antifreeze peptides on specific ice crystal surface planes. *Biophys. J.* **59**, 409–418 (1991).
11. Deng, G., Andrews, D. W. & Laursen, R. A. Amino acid sequence of a new type of antifreeze protein, from the longhorn sculpin *Myoxocephalus octodecimspinosus*. *FEBS Lett.* **402**, 17–20 (1997).
12. Wen, D. & Laursen, R. A. A model for binding of an antifreeze polypeptide to ice. *Biophys. J.* **63**, 1659–1662 (1992).
13. Knight, C. A., Driggers, E. & DeVries, A. L. Adsorption to ice of fish antifreeze glycopeptides 7 and 8. *Biophys. J.* **64**, 252–259 (1993).
14. Haymet, A. D., Ward, L. G., Harding, M. M. & Knight, C. A. Valine substituted winter flounder 'antifreeze': preservation of ice growth hysteresis. *FEBS Lett.* **430**, 301–306 (1998).
15. Zhang, W. & Laursen, R. A. Structure-function relationships in a type I antifreeze polypeptide—The role of threonine methyl and hydroxyl groups in antifreeze activity. *J. Biol. Chem.* **273**, 34806–34812 (1998).
16. Chao, H. M. *et al.* A diminished role for hydrogen bonds in antifreeze protein binding to ice. *Biochemistry* **36**, 14652–14660 (1997).
17. Graether, S. P. *et al.* Quantitative and qualitative analysis of type III antifreeze protein structure and function. *J. Biol. Chem.* **274**, 11842–11847 (1999).
18. Chou, K. C. Energy-optimized structure of antifreeze protein and its binding mechanism. *J. Mol. Biol.* **223**, 509–517 (1992).
19. Gauthier, S. Y., Kay, C. M., Sykes, B. D., Walker, V. K. & Davies, P. L. Disulfide bond mapping and structural characterization of spruce budworm antifreeze protein. *Eur. J. Biochem.* **258**, 445–453 (1998).
20. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX Pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
21. Brünger, A. T. *X-PLOR Manual Version 3.1. A system for X-ray crystallography and NMR*. (Yale Univ. Press, New Haven, Connecticut, 1992).
22. Kuszewski, J., Qin, J., Gronenborn, A. M. & Clore, G. M. The impact of direct refinement against ¹³C alpha and ¹³C beta chemical shifts on protein structure determination by NMR. *J. Magn. Reson. B* **106**, 92–96 (1995).
23. Willard, L., Wishart, D. S. & Sykes, B. D. *VADAR Version 1.2* (Univ. Alberta, Edmonton, Alberta, Canada, 1997).
24. Laskowski, R. A., MacArthur, M. W., Moss, D. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291 (1993).
25. Kraulis, P. J. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946–950 (1991).
26. Merrit, E. A. & Murphy, M. E. P. A program for photorealistic molecular graphics. Raster3D Version 2.0. *Acta Cryst. D* **50**, 869–873 (1994).
27. SYBYL Molecular modelling package. Version 6. 5 (Tripos Software, St. Louis, Missouri, 1998).

Acknowledgements

We thank S. Gauthier and D. Doucet for help with the mutagenesis, and L. Saltibus and G. McQuaid for their excellent technical assistance. Special thanks to L. Spyropoulos for helping in the analysis of the NMR data and structures. We thank Dan Garrett and the Laboratory of Chemical Physics at the National Institutes of Health for making available the program PIPP that was used in analysing our NMR data. This work was supported by MRC grants to P.L.D., B.D.S. and Z.J. and an NSERC grant to V.K.W.; Z.J. is an MRC Scholar and P.L.D. is a Killam Research Fellow.

Correspondence and requests for materials should be addressed to P. L. D. (e-mail: daviesp@post.queensu.ca). The atomic coordinates have been deposited in the Brookhaven Protein Database under accession number 1EWW.

Counting on proteomics

Download a trial of image analysis software, and seek those proteins.

ProteomeWorks

Micromass www.micromass.co.uk

Proteomics products from an alliance between BioRad and Micromass

BioRad's PDQuest image analysis software achieves rapid imaging of 2D gels, and the identification of protein spots of interest. The automated excision and enzymatic digestion of these protein spots for mass spectrometry (MS) analysis have been rate-limiting steps. The ProteomeWorks Spot Cutter is designed to eliminate the first of these two bottlenecks. Driven directly from PDQuest software, it enables automated spot excision and deposition into a microtitre plate (MTP). The excised protein spots can be robotically de-stained, chemically modified and digested in preparation for either MALDI-TOF or ESI mass spectrometry with Micromass MassPREP Station. The MassPREP Station is the robotic protein-handling system developed for The ProteomeWorks System. Incorporating Windows T based applications, the MassPREP Station includes a range of standard sample preparation and digest protocols, and user-customization of the standard protocols or the scripting of novel protocols is also supported.

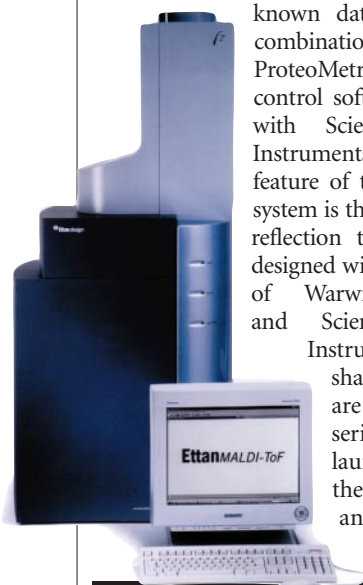
Reader Service No. 100

Etan MALDI-ToF

Amersham Pharmacia www.apbiotech.com

Adding MS to proteomics systems

The Etan MALDI-ToF mass spectrometer for protein identification was introduced earlier this year at Pittcon. The system's software matches proteins against known databases using a combination of Prowl from ProteoMetrics LLC and control software developed with Scientific Analysis Instruments Ltd. Another feature of the MALDI-ToF system is the quadratic field reflection technology (Z2) designed with Peter Derrick of Warwick University and Scientific Analysis Instruments. Amersham Pharmacia are planning a series of proteomics launches through the year, including an integrated data management



Screen proteins with Etan MALDI-ToF.

system developed with Cimarron Software to track samples through the proteomics process from 2D electrophoresis to MS.

Reader Service No. 101

MiniSpin

Eppendorf www.eppendorf.com

'Futuristic'-looking minicentrifuges

MiniSpin and MiniSpin plus microcentrifuges are suitable for general lab use. When loaded with 1.5 ml or 2.0 ml Eppendorf tubes, MiniSpin is capable of a maximum speed of 13,400 r.p.m., equivalent to 12,100g. Equipped with a black aluminium rotor, MiniSpin plus can reach 14,500 r.p.m. (14,000g). Prominent features of MiniSpin include rapid acceleration, a maximum braking time of 13 seconds, low noise levels and a footprint that is smaller than a sheet of A4 paper. MiniSpin plus has a choice between the short-spin function or continuous operation, as required, for example, for operation in a cold room.

Reader Service No. 102

SwellGel

Pierce www.piercenet.com

Protein purification made fast

SwellGel Nickel Chelating Discs bridge the gap between coated microwell plates and larger column purification. When used for performing high-throughput fusion protein purification using the 6 × His tag, the SwellGel Discs bind more than 1 mg per well, equivalent to spin column capacity in a high-throughput format. The discs are easy to use — pre-charged, pre-equilibrated and pre-dispensed. Just add clarified lysate to dried media and it is ready to purify without the need to dispense media into 96-well plates. The disc rehydrates to a nickel-chelating agarose gel within 90 seconds of adding the protein. There's no need to use expensive robotics to ensure even slurry dispensing.

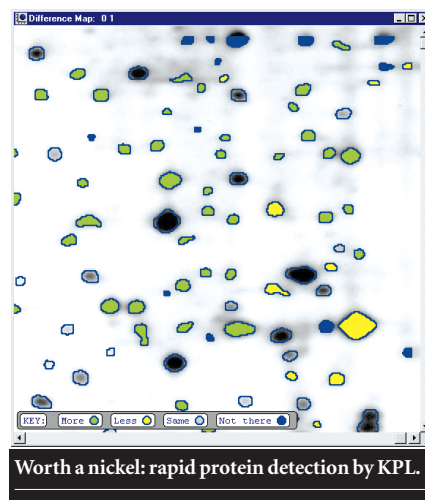
Reader Service No. 103

ExpressDetector

KPL www.kpl.com

Nickel-labelled HRP to detect His-tagged proteins

KPL's new ExpressDetector System is a simple, direct method for detecting low nanogram levels of recombinant histidine-tagged fusion proteins using nickel-labelled horseradish peroxidase. This highly specific reagent simplifies the process by eliminating the need for primary and secondary antibodies. It offers greater sensitivity on a western blot in less time than monoclonal antihistidine antibodies.



Worth a nickel: rapid protein detection by KPL.

This rapid method is recommended for use in ELISA and blotting. The reagent is offered as conjugate or in the ExpressDetector Chemiluminescent Blotting Kit, with all the components necessary for detection of his-tagged proteins on blots.

Reader Service No. 104

AlphaMatch

Alpha Innotech www.alphainnotech.com

Spot those proteins

AlphaMatch 2D image analysis software is designed for the fully automated identification, analysis and matching of protein spots in a two-dimensional gel image. The software handles molecular weight determination, density calculation, normalization, and various background subtraction methods. The user can identify proteins that share expression changes, search using a wide variety of matching criteria, and edit the results automatically or manually. AlphaMatch 2D allows importation of 8-, 12- or 16-bit tiff images from a variety of imaging sources, and the database option allows for complex queries across a great number of experiments.

Reader Service No. 105

Auto 2-D TDGS

C.B.S. Scientific www.cbsscientific.com

Automatic two-dimensional gene scanning systems

The Auto 2-D TDGS combines the efficiency of multiplex PCR with the high resolution of two-dimensional denaturing gradient gel electrophoresis. TDGS is intended for screening large numbers of individual DNA samples for mutations and polymorphisms in very large or multiple genes. It is an automated 2-D electrophoresis system that consecutively

Reader Service No. 106

Harris www.harrisphq.com

Reader Service No. 107

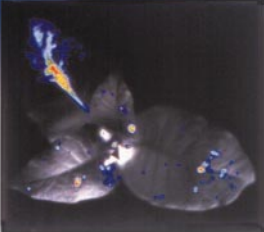
Clontech Laboratories www.clontech.com

Reader Service No. 108

Perkin-Elmer www.wallac.com

walac Advanced Imaging Systems

Image is everything



 **EG&G LIFE SCIENCES**

Reader Service No. 108

Vector Laboratories www.vectorlabs.com

ADVERTISEMENTS

a1Biotech www.a1-biotech.com

Reader Service No. 110

TotalLab www.totallab.com

Reader Service No. 111



Look no further

Ofigo Synthesis

Intuitive, easy
to use letter synthesis - easy digital synthesis of 143
English letters of up to 200 lines. This
is a unique technology and a unique feature
in your digital display.

SOON - 100% NEW - 100% NEW

**BACs • YACs • PACs
cDNAs**

Research Genetics provides clones and custom synthesis for a number of important genes and cell lines as well as cloning services. Visit our website for details.

Research Genetics
U.S. or Canada: 800-533-4763
U.K. 01880/59-1397
E.A.S. 234-526-0016

Check out our homepage at
<http://www.rgen.com>

READER ENQUIRY NO. 40